

Rapporti tecnici

INGV

**Uso dei codici di fase complementari
nei radar a compressione di impulso**

190



Direttore

Enzo Boschi

Editorial Board

Raffaele Azzaro (CT)

Sara Barsotti (PI)

Mario Castellano (NA)

Viviana Castelli (BO)

Rosa Anna Corsaro (CT)

Luigi Cucci (RM1)

Mauro Di Vito (NA)

Marcello Liotta (PA)

Simona Masina (BO)

Mario Mattia (CT)

Nicola Pagliuca (RM1)

Umberto Sciacca (RM1)

Salvatore Stramondo (CNT)

Andrea Tertulliani - Editor in Chief (RM1)

Aldo Winkler (RM2)

Gaetano Zonno (MI)

Segreteria di Redazione

Francesca Di Stefano - coordinatore

Tel. +39 06 51860068

Fax +39 06 36915617

Rossella Celi

Tel. +39 06 51860055

Fax +39 06 36915617

redazionecen@ingv.it



Rapporti tecnici INGV

USO DEI CODICI DI FASE COMPLEMENTARI NEI RADAR A COMPRESSIONE DI IMPULSO

Umberto Sciacca¹, Achille Zirizzotti²

¹INGV (Istituto Nazionale di Geofisica e Vulcanologia, Sezione Sismologia e Tettonofisica)

²INGV (Istituto Nazionale di Geofisica e Vulcanologia, Sezione Geomagnetismo, Aeronomia e Geofisica Ambientale)

190

INDICE

Introduzione	5
1. Richiami generali sui codici di fase.....	5
2. Trasmissione e ricezione del codice di fase con una sola portante	7
3. Ricezione del codice di fase usando portanti in quadratura.....	9
3.1 Estrazione del codice nel dominio del tempo	9
3.2 Estrazione del codice nel dominio della frequenza	11
4. Codici di fase complementari	13
4.1 Generalità	13
4.2 Due canali con frequenza diversa	14
4.3 Due canali con polarizzazione diversa	14
4.4 Trasmissione in sequenza nel tempo	15
4.5 Un solo canale con portanti in quadratura.....	17
5. Analisi della ricezione dei codici complementari con portanti in quadratura	17
6 Sfruttamento di due canali con polarizzazione circolare opposta	21
7. Considerazioni sulle ampiezze dei segnali ricevuti.....	23
8. Conclusioni	24
Bibliografia.....	24

Introduzione

Il metodo cosiddetto della “compressione di impulso” è noto ed usato nella tecnica radar ormai da tempo. Si tratta di un sistema di codifica del segnale trasmesso, volto ad allargare la banda di un segnale radar senza accorciare la durata dell’impulso trasmesso al fine di operare con una minore potenza di picco a parità di energia, parametro importante per non diminuire la probabilità di rivelazione dell’eco. Contemporaneamente, essendo il potere risolutivo proporzionale alla banda del segnale trasmesso, la grande banda associata all’impulso codificato rende elevato anche il potere risolutivo. Naturalmente, per poter estrarre le informazioni relative alla posizione, l’eco ricevuta deve essere opportunamente trattata, procedimento che, portando ad una riduzione della durata temporale del segnale, viene detto appunto di “compressione”.

Esistono varie tecniche di codifica, sia di tipo analogico che digitale. La codifica digitale più semplice è quella che usa “codici di fase”: la fase della portante viene modificata variandola in maniera discreta con salti di 0° o 180° a seconda del bit di codice da sovrapporre ad essa. Come contropartita dei vantaggi appena accennati, il sistema che comprime l’eco ricevuta, fa comparire, oltre al picco principale corrispondente alla posizione dell’oggetto illuminato dal radar, anche dei “lobi laterali”, dovuti al processo matematico di compressione. Tali code interferiscono con la corretta rivelazione degli echi, introducendo un “rumore” artificiale che rende più difficile l’identificazione di bersagli vicini. È per questo motivo che sono stati escogitati dei sistemi, detti a codici complementari, caratterizzati da “lobi laterali” assenti; ciò è realizzabile trasmettendo due impulsi codificati anziché uno solo, ma ciò purtroppo può comportare conseguenze non piacevoli, che vanno tenute sotto controllo.

La ionosonda AIS sviluppata in INGV [Baskaradas et al., 2002] fa uso di codici complementari trasmessi opportunamente. Nel presente lavoro, dopo alcuni richiami generali sui codici di fase, viene presentata una disamina dei possibili modi di trasmissione e ricezione dei codici complementari, compreso quello usato da AIS; in tal modo vengono messi in risalto pregi e difetti di ciascuno. Tale studio rappresenta una base di partenza per future applicazioni in altri strumenti, come il Glacio Radar [Zirizzotti et al., 2008].

1. Richiami generali sui codici di fase

Prima di parlare del motivo per cui si usano i codici complementari è bene richiamare il funzionamento dei codici di fase in generale. Quelle che saranno richiamate sono solo poche nozioni, al solo scopo di inquadrare l’argomento, peraltro ampiamente trattato in letteratura [Skolnik, 1990]. Lo schema fondamentale del sistema di codifica – decodifica è quello di fig. 1.

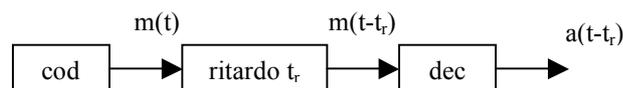


Figura 1. Schema di principio di codifica – decodifica.

Il primo blocco a sinistra (cod) è il codificatore, alla cui uscita si presenta il segnale $m(t)$. Il blocco centrale funge da modello semplificato del canale di trasmissione; si può immaginare in prima approssimazione che il segnale ricevuto arrivi identico a quello trasmesso, solamente ritardato di una quantità t_r (incognita); l’eventuale variazione di ampiezza va compensata da opportune amplificazioni, ma non porta informazioni alle quali si è interessati in questo contesto. Il segnale ricevuto $m(t-t_r)$ viene fatto passare attraverso al blocco “dec”, ottenendo un’uscita $a(t-t_r)$; se il sistema è stato progettato correttamente da quest’ultimo segnale deve essere possibile estrarre le informazioni cercate. In definitiva mentre $m(t)$ deve risultare relativamente lungo (bassa potenza di picco), $a(t)$ deve contenere un picco molto corto (alta risoluzione spaziale). Come anticipato, a seconda del codice usato, oltre al picco principale che identifica l’eco, compaiono dei “lobi laterali”, che rendono più difficoltosa l’identificazione di un oggetto quando si trovi in vicinanza di un altro. Nel seguito si parlerà solo di codici di fase binari, cioè tali che il segnale $m(t)$ assuma solo due valori o “bit” (+1 e -1 o 0 e 1), i quali andranno a variare la fase di una portante, nella maniera che sarà descritta più avanti. È dato per scontato che per trasmettere ciascun bit del codice sia necessario un certo tempo Δt , che fissa la risoluzione.

Prima di proseguire si riportano alcuni richiami matematici della teoria del passaggio di segnali nei sistemi fisici. Il decodificatore, pensato nella maniera più generale possibile, non è altro che un sistema lineare e permanente (o stazionario) che viene caratterizzato dalla sua risposta impulsiva. È noto che, se indichiamo con $h(t)$ tale risposta, l'uscita del sistema può essere scritta come segue:

$$a(t) = \int m(\tau)h(t-\tau)d\tau = m(t) * h(t). \quad (1.1)$$

Con il simbolo (*) si è indicata in maniera compatta l'operazione di convoluzione, definita dall'integrale della (1.1); si lavora nell'ipotesi che tutti i segnali siano rappresentati da funzioni reali della variabile tempo. È noto poi che, qualora si ricavano le trasformate di Fourier dei segnali presenti nella (1.1), tra essi esiste la seguente relazione, in cui si sono usate le lettere maiuscole per indicare le trasformate (es. $M(f) = \mathcal{F}\{m(t)\}$):

$$A(f) = M(f) \cdot H(f). \quad (1.2)$$

Perché l'uscita del decodificatore risulti quella corretta, esso deve "riconoscere" il segnale che gli si presenta all'ingresso; tale risultato si ha quando la sua risposta impulsiva è pari al segnale atteso ribaltato nel tempo, cioè quando:

$$h(t) = m(-t). \quad (1.3)$$

In queste condizioni infatti l'uscita viene ad essere semplicemente l'autocorrelazione del codice e la (1.1) diventa:

$$a(t) = \int m(\tau)m(t+\tau)d\tau = m(t) \circ m(t), \quad (1.4)$$

ove si è usato il simbolo (\circ) per denotare l'operazione di correlazione. La teoria infine ci ricorda che nel dominio della frequenza la correlazione diventa di nuovo un prodotto, come nella (1.2), solo che della funzione da correlare va preso il complesso coniugato (indicato con un asterisco ad esponente):

$$A(f) = M(f) \cdot H(f) = M(f) \cdot M^*(f) = |M(f)|^2. \quad (1.5)$$

Quando il segnale $m(t)$ si presenta ritardato di una quantità t_r , tale quantità comparirà nelle espressioni appena richiamate. In particolare la (1.4) diventa:

$$a'(t) = a(t-t_r) = \int m(\tau)m(t-t_r+\tau)d\tau = m(t) \circ m(t-t_r), \quad (1.6)$$

la cui trasformata vale:

$$A'(f) = A(f)e^{-j2\pi ft_r} = |M(f)|^2 e^{-j2\pi ft_r}, \quad (1.7)$$

nella (1.7) è comparso un termine di fase additivo, lineare in frequenza con pendenza data da t_r .

Tornando ora ai codici binari, un codice è costituito dalla sequenza di impulsi rettangolari e, quando si va a calcolare la correlazione tra due di essi, l'uscita avrà una forma triangolare di durata Δt , come è facile verificare; il picco del triangolo rappresenta l'ampiezza significativa. Pertanto la correlazione tra due sequenze di impulsi, $a(t)$, è data da una sequenza di triangoli, dei quali quello centrale avrà un'ampiezza maggiore, mentre prima e dopo compaiono i lobi laterali.

Tra tutti i codici possibili ve ne sono alcuni che fanno uso delle cosiddette "sequenze ottimali", che tali sono nel senso che, fissata una la lunghezza N (numero di bit), presentano il numero minimo di lobi laterali, non necessariamente tutti uguali. Sono codici ottimali i cosiddetti *codici di Barker*, i quali sono contraddistinti dalla ulteriore proprietà di produrre lobi laterali uguali. In pratica un codice di Barker di ordine N è lungo $N\Delta t$, produce all'uscita un picco principale ampio N volte i picchi dei lobi laterali ed è

lungo Δt , pertanto consente una risoluzione corrispondente a questo intervallo temporale. Il più lungo codice di Barker ha $N=13$, e presenta la seguente sequenza di bit: 1111100110101.

Come esempio di ciò che succede all'uscita del decodificatore quando una sequenza di bit viene presentata al suo ingresso, si prenda il codice con $N=5$: +1+1+1-1+1. Effettuando la correlazione di questa sequenza con una replica di sé stessa traslata, si ottiene la sequenza: 101050101. Come si vede, i lobi laterali, quando ci sono (alcuni sono nulli), sono pari ad $1/N$ del lobo principale. La lunghezza temporale di $a(t)$ è pari a $\Delta t(2N-1)$, ma il lobo principale è lungo solo Δt .

Esistono molti altri tipi di codice, ciascuno con varie proprietà e lobi laterali. Per gli scopi della presente trattazione sono sufficienti le nozioni richiamate, perché essenziali per introdurre i codici complementari, cosa che sarà fatta più avanti; prima è bene richiamare i metodi usati per realizzare la codifica e la decodifica.

2. Trasmissione e ricezione del codice di fase con una sola portante

La situazione presentata nel paragrafo precedente è particolarmente semplice, perché si è immaginato che il segnale utile viaggiasse così com'è sul canale di trasmissione. Nella pratica ciò non può accadere e, per poter avere la trasmissione a distanza, il codice deve essere sovrapposto ad una portante a radiofrequenza. Uno schema a blocchi che si avvicina maggiormente alla realtà è pertanto quello riportato in fig. 2.

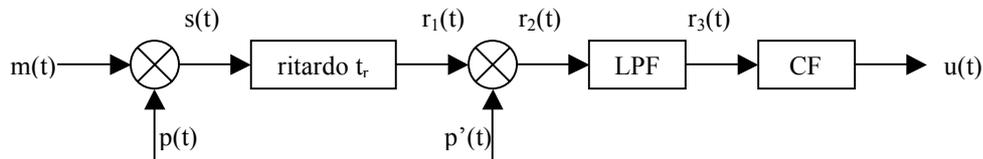


Figura 2. Schema di codifica – decodifica su portante.

Il segnale $p(t)$ è la portante, che può essere assunta uguale ad una sinusoidale, quindi il segnale trasmesso $s(t)$ non è altro che la sinusoidale modulata dal codice $m(t)$; la modulazione viene realizzata con un semplice moltiplicatore (*mixer*). L'uscita di quest'ultimo è una sinusoidale che cambia bruscamente tra valori discreti di fase ogni volta che viene trasmesso un "simbolo" del codice (nel caso di una modulazione 2PSK le fasi sono due per simbolo, quindi i possibili salti tra un simbolo e l'altro sono 0° e 180°).

Si ha per la portante ed il segnale trasmesso:

$$p(t) = \cos(2\pi f_0 t), \quad (2.1)$$

$$s(t) = m(t) \cdot \cos(2\pi f_0 t), \quad (2.2)$$

in cui si è assunto che il codice semplicemente funzioni da coefficiente moltiplicativo di ampiezza del segnale $p(t)$, la cui fase viene presa a riferimento di tutto il sistema. Tale posizione è una semplificazione che non può essere applicata in tutti i casi, dato che, a rigore, vale solo se si ipotizza una modulazione d'ampiezza. Tuttavia può essere fatta valere anche per le modulazioni di fase 2PSK, dato in esse la variazione di fase avviene tra due fasi opposte: tutto va come se la portante fosse modulata in ampiezza da un segnale che oscilla tra i due valori discreti +1 (fase "0") e -1 (fase " π "). Dato che la trattazione matematica sarà portata avanti solo per questo genere di modulazione, viene assunta valida la (2.2).

Analogamente a quanto detto nel par.1, supponiamo che il segnale trasmesso $s(t)$ arrivi sull'antenna inalterato nella forma e compaia nel segnale ricevuto $r(t)$ solamente ritardato; esso viene fatto passare attraverso il demodulatore che ha al suo ingresso di "oscillatore locale" (LO – *Local Oscillator*) il segnale $p'(t)$ (analogo a $p(t)$ a meno di una differenza di fase costante), producendo r_2 . Segue il filtro passa basso (LPF – *Low Pass Filter*) allo scopo di eliminare le componenti a frequenza doppia di quella di $p(t)$. L'uscita di quest'ultimo (r_3) viene fatta passare attraverso il "dec", d'ora in poi chiamato "filtro a correlazione" (CF – *Correlation Filter*), in modo da ottenere l'uscita e misurarne il tempo di ritardo rispetto all'ingresso. Far funzionare il sistema descritto risulta operazione piuttosto delicata, dato che l'ampiezza dell'uscita dipende

dalle relazioni di fase tra i segnali di “oscillatore locale” $p(t)$, le quali si vanno a combinare con il tempo di ritardo t_r producendo un’ampiezza in uscita che è funzione di t_r . Ciò apparirà più chiaro riportando le espressioni matematiche dei segnali introdotti.

Per il segnale ricevuto si ha:

$$r_1(t) = m(t - t_r) \cdot \cos[2\pi f_0(t - t_r)] = m(t - t_r) \cdot \cos(2\pi f_0 t - \varphi_r), \quad (2.3)$$

ove si è indicata con φ_r la quantità:

$$\varphi_r = 2\pi f_0 t_r, \quad (2.4)$$

che in generale non dipende dal tempo, ma è funzione del ritardo t_r (è vero che il ritardo con cui si presenta l’eco può cambiare, se ad esempio l’oggetto illuminato dal radar si sposta, ma ciò viene assunto essere trascurabile nel breve intervallo di tempo intercorrente tra due impulsi). Ora calcoliamo l’uscita del demodulatore; il segnale di oscillatore locale $p'(t)$ in generale sarà caratterizzato da una differenza di fase φ_0 rispetto a $p(t)$ (usando le formule di Werner):

$$p'(t) = \cos(2\pi f_0 t + \varphi_0), \quad (2.5)$$

$$\begin{aligned} r_2(t) &= m(t - t_r) \cdot \cos(2\pi f_0 t - \varphi_r) \cdot \cos(2\pi f_0 t + \varphi_0) = \\ &= \frac{m(t - t_r)}{2} \cdot [\cos(4\pi f_0 t - \varphi_r + \varphi_0) + \cos(\varphi_0 + \varphi_r)] \end{aligned} \quad (2.6)$$

Dopo il filtraggio, la componente a frequenza doppia $2f_0$ viene eliminata, rimanendo solo l’altra, in “banda base”:

$$r_3(t) = \frac{m(t - t_r)}{2} \cdot \cos(\varphi_r + \varphi_0) = \frac{m(t - t_r)}{2} \cdot \cos(2\pi f_0 t_r + \varphi_0). \quad (2.7)$$

Ora, le informazioni sul ritardo sono contenute nel segnale $m(t - t_r)$, pertanto quando si fa passare nel CF il segnale $r_3(t)$, l’uscita prodotta avrà la forma di $a(t)$, solo ritardata. Purtroppo è presente il termine di ampiezza $\cos(2\pi f_0 t_r + \varphi_0)$, che può variare a seconda della distanza dell’oggetto bersaglio, ma che comunque è sempre minore di uno. Un modo per evitare tale perdita di ampiezza (che poi si evidenzia con una diminuzione della probabilità di rivelazione) è quello di non usare una demodulazione coerente (“omodina”) come quella di fig. 2, ma una ad “involuppo”. Alternativamente si può utilizzare un sistema con “portanti in quadratura”, che verrà descritto tra breve.

In definitiva, dato che il segnale p' viene generato assieme a p , è possibile far assumere ad esso un valore di fase prefissato, rendendo p e p' perfettamente sincroni. Per semplificare i calcoli poniamo:

$$\varphi_0 = 0^\circ, \quad (2.8)$$

cosicché si abbia:

$$r_3(t) = \frac{m(t - t_r)}{2} \cdot \cos \varphi_r; \quad (2.9)$$

l’uscita è quindi:

$$u(t) = \frac{\cos \varphi_r}{2} \cdot m(t - t_r) \circ c(t) = \frac{\cos \varphi_r}{2} \cdot a(t - t_r), \quad (2.10)$$

in cui la funzione $a(t)$ è la stessa di fig. 1; essa risulta pertanto traslata di t_r e alterata in ampiezza di $\cos \varphi_r$.

3. Ricezione del codice di fase usando portanti in quadratura

3.1 Estrazione del codice nel dominio del tempo

Al fine di evitare la perdita di ampiezza dovuta a $\cos\varphi_r$, tutto ciò che in fig.2 è a valle di $r_1(t)$ può essere sostituito dallo schema di fig. 3 (suggerito da [Skolnik, 1990]).

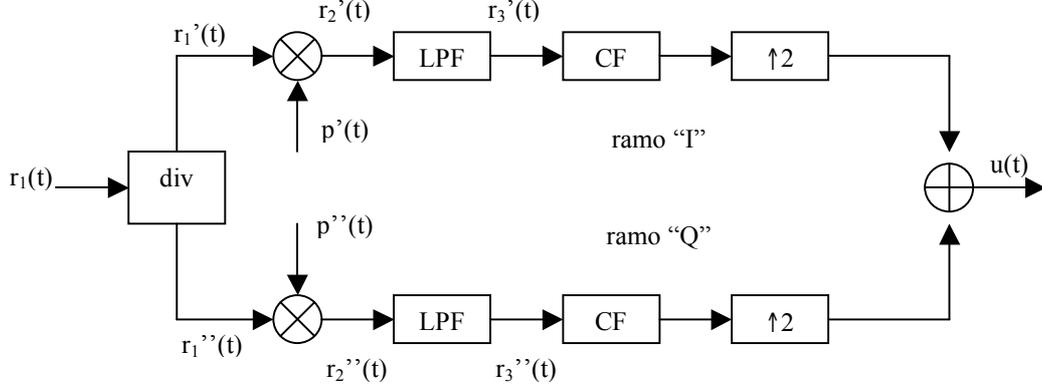


Figura 3. Demodulazione con portanti in quadratura nel dominio del tempo.

In essa i due segnali di OL sono in quadratura di fase, cioè sono identici ma con una differenza di fase di 90° , questo è il motivo dei nomi (del tutto convenzionali) che in genere si danno ai due rami del ricevitore: I (*In phase*) e Q (*Quadrature*) rispetto alla fase della portante. Il blocco “div” è un separatore, che suddivide l’ingresso in due segnali di uguale energia; pertanto, idealmente, si ha:

$$r_1'(t) = r_1''(t) = \frac{r_1(t)}{\sqrt{2}}; \quad (3.1)$$

mentre per le portanti di OL si ha (ricordare la posizione fatta per la fase):

$$\begin{aligned} p'(t) &= \cos(2\pi f_0 t) \\ p''(t) &= \cos(2\pi f_0 t + \pi/2) \end{aligned} \quad (3.2)$$

Svolgendo i calcoli per il ramo I si trova:

$$r_2' = r_1' \cdot p' = \frac{m(t-t_r)}{\sqrt{2}} \cdot \cos(2\pi f_0 t + \varphi_r) \cdot \cos(2\pi f_0 t) = \frac{m(t-t_r)}{2\sqrt{2}} \cdot [\cos(4\pi f_0 t + \varphi_r) + \cos(\varphi_r)], \quad (3.3)$$

$$r_3'(t) = \frac{m(t-t_r)}{2\sqrt{2}} \cdot \cos\varphi_r. \quad (3.4)$$

Analogamente, per il ramo Q si trova:

$$r_3''(t) = \frac{m(t-t_r)}{2\sqrt{2}} \cdot \sin\varphi_r. \quad (3.5)$$

Pertanto, quando si fa passare il segnale attraverso il CF e poi si fa il quadrato, si trova:

$$u(t) = \left[\frac{\cos\varphi_r}{2\sqrt{2}} \cdot a(t-t_r) \right]^2 + \left[\frac{\sin\varphi_r}{2\sqrt{2}} \cdot a(t-t_r) \right]^2 = \frac{a^2(t-t_r)}{8}. \quad (3.6)$$

Quanto esposto può essere espresso in forma più compatta se si pensa alla combinazione dei segnali r_3' e r_3'' come un unico segnale complesso che può essere espresso nella forma che evidenzia un modulo ed una fase:

$$r_3(t) = r_3'(t) + j \cdot r_3''(t) = \frac{m(t-t_r)}{2\sqrt{2}} \cdot (\cos\varphi_r + j \sin\varphi_r) = \frac{m(t-t_r)}{2\sqrt{2}} \cdot e^{j\varphi_r}, \quad (3.7)$$

$$u(t) = |r_3(t) \circ m(t)|^2 = \left| \frac{a(t-t_r)}{2\sqrt{2}} \cdot e^{j\varphi_r} \right|^2 = \frac{a^2(t-t_r)}{8}$$

A rigore l'uscita (3.6) andrebbe poi sottoposta a radice quadrata, ma le distorsioni di ampiezza che conseguono al fatto che è presente un quadrato non sono significative, dato che l'essenziale è che venga preservata la forma del segnale dopo compressione, con il lobo principale e quelli laterali; pertanto l'operazione di quadratura non ha conseguenze da questo punto di vista. La (3.6) rispetto alla (2.10) è alterata solo di un fattore di ampiezza poco significativo (su di esso si ritornerà più avanti); la cosa più importante è che è sparita la dipendenza da φ_r .

L'elevazione al quadrato però potrebbe portare ad un diverso problema: si è supposto che il codice ricostruito avesse solo lobi a valori positivi (tale è ad es. il codice di Barker); qualora esso avesse anche valori negativi, essi andrebbero persi dopo la quadratura. Come si vedrà tra breve, può accadere che i segni della funzione in uscita dal filtro a correlazione siano importanti al fine dell'esecuzione corretta di operazioni successive, pertanto, in quei casi non ci si può accontentare della semplice elevazione al quadrato di fig. 3.

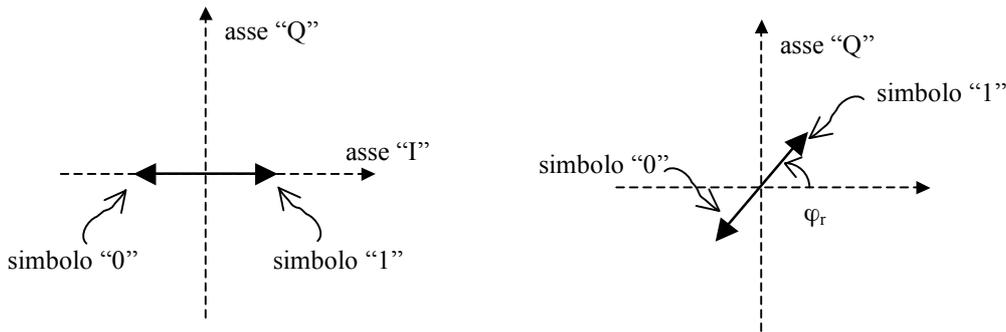


Figura 4. Diagrammi vettoriali dei codici in trasmissione e ricezione.

Per capire meglio i motivi di quanto espresso dalle (2.10) e (3.6), si osservino i diagrammi vettoriali fig. 4. Tali diagrammi sono delle rappresentazioni semplificate di quanto accade alla fase e all'ampiezza di una sinusoidale, rappresentati sul piano complesso da modulo e argomento di un vettore (talvolta tale piano viene detto piano dei fasori). Si suppone che tutte le sinusoidi da rappresentare abbiano la stessa frequenza, che può essere quella della portante effettiva o una sua traslazione (come ad esempio quella di una media frequenza di un ricevitore, che non è stata mai rappresentata per semplicità). Quindi gli unici parametri che possono variare sono l'ampiezza e la fase; tralasciamo al solito le variazioni di ampiezza e concentriamoci su quelle di fase. La sinusoidale in trasmissione può assumere solo le fasi a 0° o 180° , cosa rappresentata dalle due frecce nel diagramma di sinistra. Quando l'onda torna al ricevitore, arriva con una fase φ_r , quindi i vettori rappresentativi appaiono come nel diagramma di destra: anche se il modulo non si è alterato, ogni volta che si tenta di demodulare si perde una parte dell'ampiezza; la demodulazione infatti corrisponde a prendere la proiezione sull'asse "reale" (se si moltiplica per un coseno) o sull'asse "immaginario" (se si moltiplica per un seno), o su qualunque altra retta (moltiplicazione per un coseno con una fase aggiunta arbitraria). Pur attenuate in intensità tali proiezioni mantengono le informazioni sul segno ed il passaggio attraverso il filtro adattato non le altera; invece esse si perdono quando eleviamo al quadrato. Per fortuna il filtraggio adattato viene effettuato prima della quadratura, pertanto se il segno non serve per altri motivi (come si vedrà per i codici complementari), il sistema di fig. 3 riesce a recuperare tutta l'ampiezza.

L'estrazione dell'ampiezza può essere effettuata anche in altri modi. In [Cabrera et al., 2010] si trova un confronto tra lo schema di fig. 3 (ivi chiamato QCC) ed uno schema alternativo; in quel lavoro si prende in considerazione un sistema più complesso di quello di figura 3, dato che viene adottato un codice

complementare e un sistema di integrazione di più impulsi ricevuti. In ogni caso il problema della eliminazione della dipendenza da φ_r senza alterare le informazioni sul segno può essere risolto solo andando a lavorare direttamente sul vettore nel piano complesso, elaborando in modo separato l'informazione contenuta nello sfasamento, che equivale al tempo di ritardo.

3.2 Estrazione del codice nel dominio della frequenza

Data la perfetta corrispondenza tra i due domini del tempo e della frequenza, dovrebbe essere possibile ottenere lo stesso risultato del sistema di figura 3, che opera esclusivamente nel tempo, adottandone un altro che operi nel dominio della frequenza. In realtà svolgere le operazioni nel dominio della frequenza, oltre a semplificare l'operazione di correlazione (che diventa un semplice prodotto), ci mette in grado di intravedere un modo per eliminare la dipendenza da φ_r .

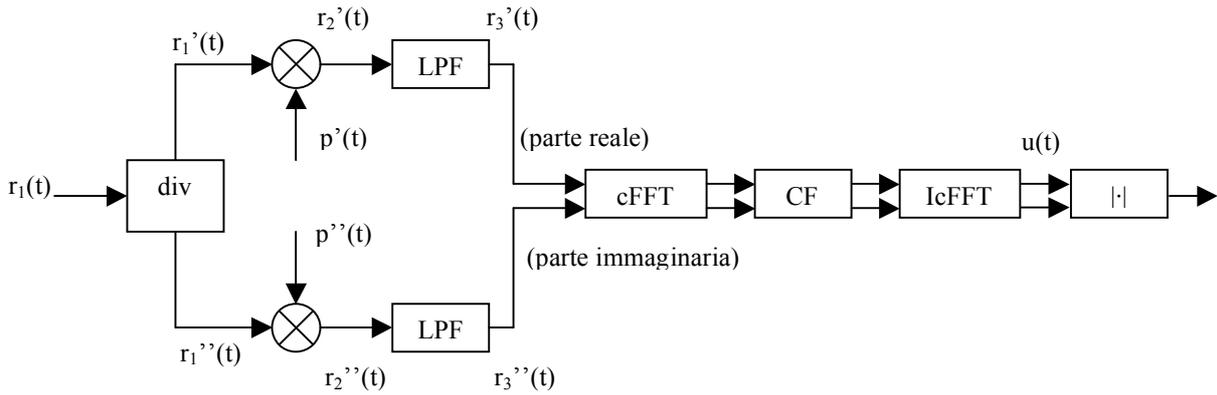


Figura 5a. Demodulazione con portanti in quadratura nel dominio della frequenza.

Il nuovo sistema di elaborazione è riportato in fig. 5a. La parte a monte, fino ai filtri LPF è analoga a quella di fig. 3, solo che ora le due uscite vengono pensate come parte reale e parte immaginaria di un unico segnale complesso, come già visto con la (3.7). Esso viene dapprima trasformato (blocco “cFFT” cioè “complex FFT”), quindi viene effettuata la correlazione ed infine riportato il tutto nel dominio del tempo (blocco “IcFFT” dove “I” sta per “inverse”). Alla fine viene calcolato il modulo del segnale, dato che in generale non è detto che l'uscita della IFFT sia un segnale reale, quindi bisogna riottenere uno a partire dai due disponibili.

Che le cose vadano proprio come atteso si può verificare facilmente. Si riprendano infatti i segnali dati dalle (3.4) e (3.5). Il segnale complesso da far elaborare all' algoritmo FFT è dato dalla (3.7).

Il calcolo della trasformata di Fourier di un tale segnale non è immediato, però è possibile arrivarci ricordando alcune proprietà delle trasformate. Usando le lettere maiuscole per denotare le trasformate dei segnali (es. $m(t)$ trasformato diventa $M(f)$), ponendo $\varphi_r = 2\pi f_0 t_r$ e usando la pulsazione ω per evitare di riportare tutte le volte il fattore 2π ($2\pi f = \omega$), si hanno i seguenti passaggi intermedi:

$$F\{m(t - t_r)\} = M(\omega)e^{-j\omega t_r} = |M|e^{j\text{Arg}(M)}e^{-j\omega t_r}, \quad (3.8)$$

$$R_3(\omega) = \frac{|M(\omega)|}{2\sqrt{2}} \cdot (\cos \varphi_r + j \sin \varphi_r) \cdot e^{j\text{Arg}(M)} e^{-j\omega t_r}. \quad (3.9)$$

La trasformata dell'uscita quindi è (ricordare la 1.5):

$$U(\omega) = R_3(\omega) \cdot M^*(\omega) = \frac{|M(\omega)|^2}{2\sqrt{2}} \cdot (\cos \varphi_r + j \sin \varphi_r) \cdot e^{-j\omega t_r} = \frac{|M(\omega)|^2}{2\sqrt{2}} \cdot e^{j\varphi_r} \cdot e^{-j\omega t_r}. \quad (3.10)$$

La (3.10) è formalmente analoga alla (1.7) che esprime l'uscita ideale: il termine esponenziale a terzo membro contiene un contributo che rappresenta una fase linearmente variabile con la frequenza e pendenza t_r

(cioè l'informazione rilevante sul ritardo); vi è poi un altro contributo, costante ($j\varphi_r$), che rappresenta il "disturbo" dovuto alla rotazione di fase della portante.

L'analisi della (3.10) porta alla seguente considerazione: se effettuiamo l'anti trasformata senza ulteriori elaborazioni, ritroviamo i due segnali che erano presenti nei rami I e Q della fig.3 prima della elevazione a quadrato. Per rendersene conto si immagini che sia $\varphi_r=0$; in tal caso la (3.10) ha un modulo che è funzione pari della frequenza ed una fase funzione dispari (è data dal solo termine lineare); ciò corrisponde ad un segnale nel dominio del tempo puramente reale. In pratica sui due rami di uscita della IcFFT si ritrova un segnale nel ramo superiore (parte reale) e zero in quello inferiore (a parte il rumore, ma stiamo considerando un caso ideale senza rumore). Questo caso corrisponde ad avere una rotazione nulla dei fasori di fig.4. La condizione $\varphi_r \neq 0$ corrisponde ad una semplice rotazione dei fasori, senza variazione dell'ampiezza; all'uscita compare un segnale (posizionato allo stesso ritardo t_r) anche nel ramo inferiore della IcFFT. Ciò, in termini matematici, equivale a pensare un'uscita con parte immaginaria non nulla, infatti l'aggiunta nella (3.10) di una fase costante spezza la simmetria dispari della fase che si aveva per $\varphi_r=0$. Per riottenere un solo segnale in uscita basterebbe estrarre il modulo; tutto questo però sembra non aver portato ad alcun giovamento rispetto allo schema di fig.3. Diverso sarebbe il caso se si riuscisse ad eliminare il termine di fase, cosa che, almeno in via di principio, non è difficile, aiutandosi con la fig.5b.

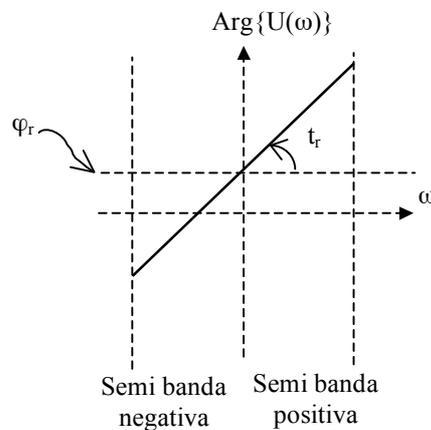


Figura 5b. Eliminazione della rotazione di fase della portante.

Nella figura è chiaramente visibile il termine costante φ_r ; esso può essere ricavato facilmente dal valore della funzione $Arg\{U\}$ per $f=0$. Se non si dispone della funzione di fase in forma esplicita, ma solo delle parti reale ed immaginaria, il valore cercato si può lo stesso ricavare dalla:

$$\varphi_r = \arctan\left(\frac{\text{Im}\{U(0)\}}{\text{Re}\{U(0)\}}\right). \quad (3.11)$$

Una volta noto φ_r lo si può sottrarre dalla (3.10) operando in questo modo la rotazione sui fasori, riportandoli a coincidere con l'asse reale di fig.4. Nel calcolo della (3.11) va solo posta qualche attenzione al fatto che l'arco tangente tende a riportare l'angolo sempre al I e IV quadrante, cosa che con opportuni accorgimenti si può correggere.

Purtroppo il risultato dello sforzo teso ad eliminare il termine di fase costante nella (3.10), pur elegante, può essere applicato solo nel caso in cui esista un solo segnale di eco, come si è sempre implicitamente ammesso finora. Andando a considerare il caso in cui siano presenti in ricezione più echi, si ha che ciascuno porta con sé un termine di fase φ_r differente, perché i ritardi sono diversi e quindi lo sono anche gli sfasamenti con cui si presentano le portanti sottostanti. Ora, nel dominio della frequenza tutti gli echi hanno spettri sovrapposti e non c'è modo di separare i termini di fase di ciascuno. Pertanto, l'unica maniera per recuperare l'ampiezza massima è tornare nel dominio del tempo ed estrarre il modulo; gli echi saranno separati sulla base del tempo di arrivo (ammesso che la banda del sistema consenta di risolverli).

Anche se apparentemente non porta vantaggi rispetto al sistema operante nel dominio del tempo, si è ritenuto di esporre il funzionamento anche di quello nel dominio della frequenza per completare il quadro

concettuale. In realtà operare nel dominio della frequenza rimane preferibile poiché determina vantaggi di altra natura (come la facilità di calcolo della correlazione e di filtraggio di frequenze indesiderate).

4. Codici di fase complementari

4.1 Generalità

Come si è detto, l'uso dei codici di fase consente un notevole miglioramento delle prestazioni: a parità di potenza di picco trasmessa, l'energia dell'impulso aumenta in ragione del numero di elementi del codice N , consentendo una migliore rivelazione in presenza di rumore, mantenendo la risoluzione a valori corrispondenti alla durata di un elemento di codice Δt . Come visto nel paragrafo iniziale, in generale il segnale di uscita dal filtro adattato $a(t)$ porta con sé i lobi laterali, che agiscono da disturbo, motivo per cui si è cercato un sistema che eliminasse questo fastidioso inconveniente; tale metodo fa uso delle cosiddette "sequenze complementari". Si tratta di due sequenze di codice che forniscono due uscite $a(t)$ del tutto analoghe, con il lobo principale identico, ma con i lobi laterali di segno opposto; in tal modo, sommando le due uscite, è possibile annullare completamente i lobi laterali (almeno idealmente, in pratica la cancellazione non è mai perfetta).

Come esempio del funzionamento di una coppia di tali sequenze, si prendano le due seguenti, con $N=4$:

$$\begin{array}{rcccc} \text{seq.A:} & +1 & +1 & +1 & -1 \\ \text{seq.B:} & +1 & +1 & -1 & +1 \end{array}$$

Calcolando le funzioni di autocorrelazione di entrambe, si ottengono le sequenze seguenti, che sommate, danno lobi laterali nulli (si noti che ciascuna sequenza ha i lobi laterali disuguali):

$$\begin{array}{rcccccc} \text{corr.A:} & -1 & 0 & +1 & +4 & +1 & 0 & -1 \\ \text{corr.B:} & +1 & 0 & -1 & +4 & -1 & 0 & +1 \\ \text{somma:} & 0 & 0 & 0 & +8 & 0 & 0 & 0 \end{array}$$

È il caso di evidenziare cosa succede ai segnali quando si usano codici complementari. Le relazioni introdotte nel par.1 devono continuare a valere per ciascuno dei due codici complementari, che possiamo chiamare m_1 e m_2 . Le (1.6) e (1.7) pertanto si sdoppiano e diventano:

$$a_i'(t) = a_i(t - t_r) = \int m_i(\tau) m_i(t - t_r + \tau) d\tau = m_i(t) \circ m_i(t - t_r), \text{ con } i=1,2, \quad (4.1)$$

$$A_i'(f) = A_i(f) e^{j2\pi f t_r} = |M_i(f)|^2 e^{-j2\pi f t_r}; \quad (4.2)$$

ma all'uscita si deve avere:

$$a''(t) = a_1(t - t_r) + a_2(t - t_r), \quad (4.3)$$

$$A''(f) = [A_1(f) + A_2(f)] \cdot e^{j2\pi f t_r} = \{|M_1(f)|^2 + |M_2(f)|^2\} \cdot e^{-j2\pi f t_r}. \quad (4.4)$$

La (4.3) esprime il fatto che, se i codici sono stati ricevuti correttamente, la funzione in uscita a'' deve avere lobi laterali nulli; analogamente nel dominio della frequenza alla somma dei moduli M_1+M_2 corrisponde un segnale nel tempo con lobi laterali nulli. Tutto lo studio che segue è dunque teso a trovare realizzazioni pratiche del metodo qui esposto in generale. Rispetto all'uso di una sola sequenza, ci si aspetta una maggiore complessità del sistema trasmettitore – ricevitore, per il fatto che esso dovrà realizzare il modo di trasmettere e ricevere i due codici su due "canali" differenti; nello stesso tempo bisogna evitare che essi interferiscano tra di loro o arrivino con ritardi t_r differenti, cosa che non renderebbe più valide le (4.3) e (4.4).

4.2 Due canali con frequenza diversa

Un primo modo di trasmettere i due codici fa uso di due canali separabili sulla base della occupazione spettrale, cioè si fa uso diverse frequenze portanti. Esse devono essere tali da poter essere facilmente separate al momento della ricezione da opportuni filtri. Si tratta di un sistema concettualmente semplice, che trasmette i due codici contemporaneamente, con lo svantaggio che porta ad un *hardware* sostanzialmente raddoppiato. La differente frequenza comporta il rischio di differenti modalità di propagazione, con complicazioni al momento della ricostruzione della sequenza (se i due echi non sono uguali la cancellazione dei lobi laterali non è perfetta). Per evitare che la differente frequenza comporti differenti caratteristiche di propagazione è opportuno usare frequenze quanto più vicine possibile, con la conseguenza positiva della possibilità di usare un'unica antenna; la contropartita è che, se si avvicinano troppo le bande, c'è il rischio di una loro non perfetta separazione in ricezione, con filtri più difficili da realizzare.

Lo schema di principio per realizzare un sistema del genere è riportato in fig. 6. Nella parte alta è rappresentato il trasmettitore: i due codici vanno a modulare due portanti a frequenza diversa (f_{R1} e f_{R2}), i successivi due filtri passa banda assicurano che in ciascuna delle loro uscite non vi siano residui che possano ricadere nella banda del canale complementare. Nella parte bassa della figura vi è il ricevitore, per il quale, a differenza delle figure precedenti, sono riportati anche gli stadi a media frequenza, dato che sono fondamentali per separare i due canali. Il segnale ricevuto infatti viene suddiviso in due canali e fatto battere con l'oscillatore locale a frequenza f_L . I successivi filtri hanno la stessa larghezza di banda di quelli in trasmissione, ma sono centrati sulle frequenze opportune: BPF₃ a $f_{R1}-f_L$, BPF₄ a $f_{R2}-f_L$. In questo modo i segnali alle loro uscite contengono solo un codice, pertanto i blocchi successivi non sono altro che quelli già incontrati, descritti nelle fig. 3 o 5a, solo che i correlatori in essi contenuti ovviamente effettueranno le correlazioni con codici diversi. L'implementazione di questo metodo è semplice e non vengono riportati sviluppi analitici.

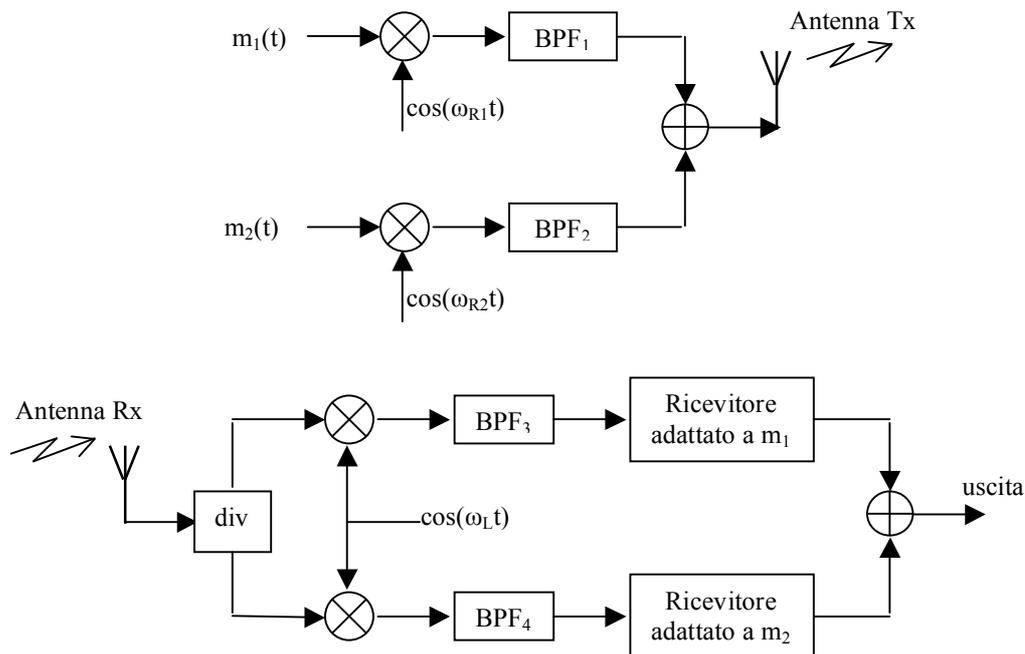


Figura 6. Sistema trasmittente e ricevente per codici complementari a “divisione di frequenza”.

4.3 Due canali con polarizzazione diversa

Un secondo metodo di trasmissione contemporanea fa uso di una differente polarizzazione lineare: se le due portanti vengono inviate a due antenne a polarizzazione lineare ortogonali, i due segnali dovrebbero viaggiare in maniera del tutto separata ed essere ricevuti da altrettante antenne ortogonali.

In fig. 7 è rappresentato lo schema di principio di un sistema basato su questo metodo. Una soluzione del genere presuppone che i segnali a diversa polarizzazione viaggino allo stesso modo; se da un lato i mezzi di propagazione sono quasi sempre isotropi, pertanto le onde a diversa polarizzazione viaggiano con analoghe

modalità, dall'altro lato può accadere che le due onde non mantengano il loro piano di polarizzazione perfettamente orientato; come conseguenza ogni canale riceve un residuo dell'altro. Anche in questo caso quindi vi è il rischio di una non perfetta indipendenza dei due canali. Per ovviare a questo inconveniente si può pensare ad un sistema che sfrutti la polarizzazione circolare; di esso si parlerà estesamente in un successivo paragrafo.

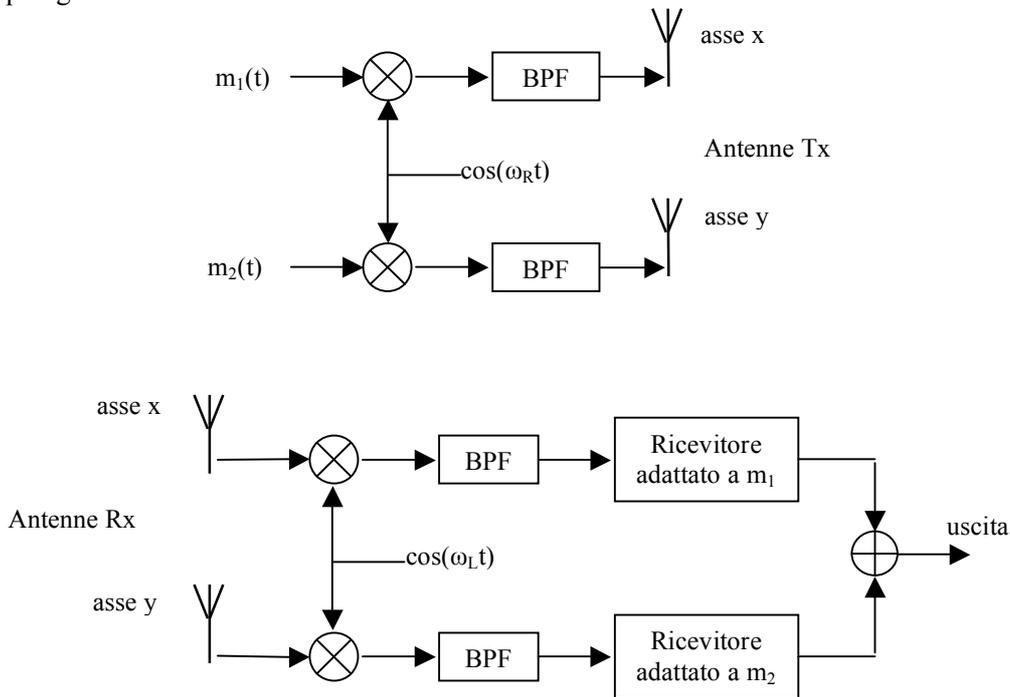


Figura 7. Sistema trasmittente e ricevente per codici complementari a “divisione di polarizzazione”.

4.4 Trasmissione in sequenza nel tempo

Un terzo sistema di trasmissione può essere chiamato a “divisione di tempo”: in pratica si trasmettono e si ricevono i due codici separati nel tempo, in sequenza. Si attende il “tempo di volo” relativo al primo impulso, codificato con il primo codice, memorizzando quanto ricevuto; poi si trasmette il secondo codice, si attende anche il secondo “tempo di volo” e si sommano i due echi. Il sistema è virtualmente esente dai problemi visti per i precedenti, e anche la complessità circuitale è inferiore rispetto ad una completa duplicazione di tutto l'hardware. La limitazione di questo sistema risiede nell'aver ipotizzato che l'oggetto illuminato dal radar non si muova nel breve tempo tra una trasmissione e la successiva; in caso contrario quanto memorizzato in seguito alla prima trasmissione andrà ad occupare celle di memoria differenti rispetto a quelle della seconda (indipendentemente dal sistema usato per la memorizzazione, ad ogni tempo di ritardo deve essere associata una posizione in cui memorizzare l'intensità dell'eco). Nella fig. 8 è riportato uno schema di quanto accade nel caso visto in 4.1 con $N=4$, nell'ipotesi che una singola cella di memoria temporale coincida con il tempo Δt e che la seconda eco arrivi con un ritardo aggiuntivo proprio pari a Δt .

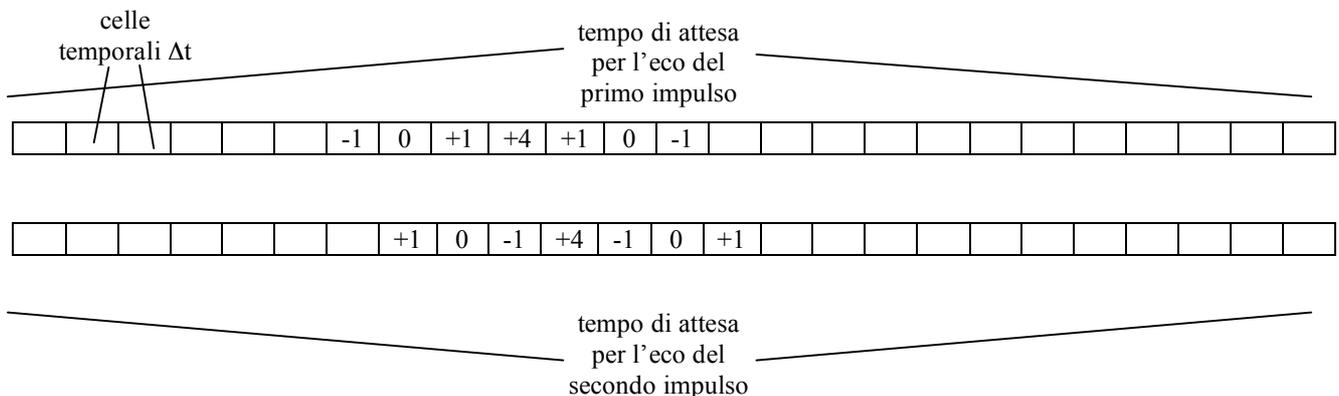


Figura 8. Schema della ricezione di due codici non allineati.

Il risultato della ricezione è: -1, 1, 1, 3, 5, -1, -1, 1, ben lontano dall'annullamento dei lobi laterali. Si noti che si è preso in esame il caso in cui lo scarto nei ritardi coincidesse esattamente con un elemento Δt . Nei sistemi reali la cella di acquisizione in distanza può essere inferiore a Δt , ed anche la differenza nei tempi di arrivo può essere inferiore a Δt ; in questi casi l'effetto è meno visibile, ma può risultare sempre fastidioso.

Ci si può chiedere qual è il limite di utilizzo di un tale sistema, che è evidentemente funzione di alcuni parametri di progetto del radar. In primo luogo esso dipenderà dalla risoluzione che si intende ottenere, da realizzare con opportuna scelta della durata del singolo elemento di codice, che in precedenza è stata chiamata Δt ; in secondo luogo dipenderà dalla distanza massima alla quale ci si aspetta di incontrare oggetti da illuminare, infatti prima di trasmettere il secondo impulso della coppia, bisognerà attendere il tempo necessario perché il primo abbia coperto tale distanza massima e sia tornato indietro, quindi, se chiamiamo H tale "altezza" massima, e c la velocità della luce (da sostituire dal valore effettivo del mezzo, se non è il vuoto), il tempo di attesa ΔT è dato da:

$$\Delta T = \frac{2H}{c}.$$

Ora, se l'oggetto si muove a velocità v rispetto al radar, nel tempo ΔT lo spazio percorso a quella velocità deve essere molto inferiore a quello che compete alla cella di risoluzione in distanza δh :

$$\delta h = \frac{c \cdot \Delta t}{2};$$

pertanto si può scrivere che deve essere:

$$v \ll \frac{\delta h}{\Delta T} = \frac{c \cdot \Delta t}{2} \cdot \frac{c}{2H} = \frac{\Delta t \cdot c^2}{4H}.$$

Dato che i riflettori ionosferici si possono considerare fermi, facciamo un esempio in un caso in cui esiste una velocità relativa tra riflettore e sistema radar. In un radar per prospezioni glaciologiche, la distanza tra gli impulsi è di circa $\Delta T = 10ms$; anche considerando un'altezza H di 1500 km (sufficiente anche per altri usi geofisici) e ipotizzando una $\Delta t = 125ns$ (che corrisponde ad una risoluzione spaziale $\delta h = 18m$), si ottiene come valore limite per la velocità v circa 1870 m/s, cosa che fa vedere come la possibilità di una non perfetta cancellazione dei lobi laterali non costituisca un problema anche per questo genere di applicazioni.

Le considerazioni precedenti hanno fatto propendere per l'adozione di questo sistema per la ionosonda AIS – INGV, infatti le velocità di spostamento dei riflettori posti nella ionosfera sono estremamente piccole. Lo schema a blocchi del sistema è simile a quello di fig.5a, ove però all'uscita del correlatore nel dominio della frequenza vi è un registro che memorizza temporaneamente la ricezione del primo codice. Si sarebbe potuto usare anche lo schema di fig.3, con due registri di memoria posti prima dei due quadratori: la somma dei due codici complementari, pur in condizioni di ampiezza non massima, è esente da lobi laterali, pertanto i successivi elevamenti al quadrato portano al recupero dell'ampiezza senza perdite di informazione.

Dato che AIS-INGV opera nel dominio della frequenza, quando viene ricevuto il secondo codice, la somma dei due spettri dà luogo ad un segnale antitrasformato nel tempo senza lobi laterali; infatti, in analogia alla (3.10), gli spettri relativi ai due codici avranno la forma seguente:

$$U_1(\omega) = \frac{|M_1(\omega)|^2}{2\sqrt{2}} \cdot e^{j\varphi_r} \cdot e^{-j\omega t_r}, \quad (4.5)$$

$$U_2(\omega) = \frac{|M_2(\omega)|^2}{2\sqrt{2}} \cdot e^{j\varphi_r} \cdot e^{-j\omega t_r}. \quad (4.6)$$

L'ipotesi è che, nel breve intervallo tra i due impulsi codificati con i due codici, non vari né t_r né φ_r . Pertanto la somma fornisce:

$$U_1(\omega) + U_2(\omega) = \frac{|M_1(\omega)|^2 + |M_2(\omega)|^2}{2\sqrt{2}} \cdot e^{j\varphi_r} \cdot e^{-j\omega t_r} = \frac{|M_0(\omega)|}{2\sqrt{2}} \cdot e^{j\varphi_r} \cdot e^{-j\omega t_r}, \quad (4.7)$$

ove si è indicato con M_0 il risultato della somma di M_1 e M_2 che, ricordando la (4.4) è esente da lobi laterali. A questo punto non è più necessario adottare sistemi di recupero della fase come quello di fig.5b: semplicemente all'uscita, una volta riottenuto il segnale nel tempo, dato che esso è complesso, basta estrarre il modulo. Non si riportano ulteriori approfondimenti su questa realizzazione perché sono reperibili nelle pubblicazioni citate in bibliografia, in particolare [Baskaradas et al., 2002] e [Cabrera et al., 2010].

4.5 Un solo canale con portanti in quadratura

La presenza di due bande laterali in ogni modulazione AM, che sono identiche o con relazioni ben determinate tra esse, fa già supporre che vi sia ridondanza nella trasmissione, ed infatti in taluni casi viene usata la modulazione a bande laterali indipendenti per raddoppiare la capacità del canale a parità di banda usata nello spettro elettromagnetico.

Tale ridondanza si può sfruttare in maniera differente trasmettendo due segnali distinti su due portanti in quadratura. Questo caso è peculiare: non si sfruttano due canali distinti ma un unico canale in cui le informazioni possono rimanere inalterate perché le portanti in quadratura di fase costituiscono due "mondi" che non si influenzano (domini ortogonali), purché si disponga di un sistema analogo in ricezione che sia perfettamente agganciato in fase alle portanti ricevute. Ciò, nel caso del radar, costituisce un problema.

Quando il sistema delle portanti in quadratura è usato nei normali sistemi di telecomunicazione (ad esempio per portare l'informazione di colore nel sistema PAL della televisione analogica), assieme al segnale viene trasmesso anche qualcosa che consenta al ricevitore di ricostruire la portante; ad esempio, nella menzionata TV, prima dell'informazione relativa ad ogni riga viene mandato un piccolo treno d'onde perfettamente sincrono con la portante di colore; il ricevitore pertanto può agganciare il proprio oscillatore locale a questo treno ed effettuare la demodulazione. Il tutto presuppone che i segnali arrivino sufficientemente intensi da consentire la rivelazione senza difficoltà.

Nel caso dei sistemi radar ci si trova in condizioni molto peggiori e capita spesso di trovarsi di fronte ad oggetti che rinviano indietro echi con scarsissima energia. Per poter identificare correttamente gli echi il ricevitore dovrebbe estrarre una portante per ogni eco codificata, infatti, dato che ogni eco è contraddistinta da una distanza diversa, la fase con cui viene ricevuta la portante relativa è differente. Ora, ciò che spinge ad usare i codici, complementari o no che siano, è proprio il fatto che molti segnali sono al limite della rivelazione, mentre la presenza del codice fornisce un surplus di intensità che consente al segnale di emergere dal rumore; una portante non modulata pertanto rischierebbe di non essere ricevuta o di non essere di intensità sufficiente a far agganciare ad essa l'oscillatore locale del ricevitore.

In definitiva, in assenza dell'informazione di fase della portante di ciascun eco ricevuta, questo sistema non è in grado di mantenere separate le informazioni sui due codici. Tale circostanza, pur negativa, merita un approfondimento, pertanto nel prossimo paragrafo verrà analizzata in dettaglio.

5. Analisi della ricezione dei codici complementari con portanti in quadratura

Ipotizziamo che in trasmissione si vogliono usare due codici $m_1(t)$ e $m_2(t)$; essi possono essere di tipo qualunque, ma nel nostro caso supporremo che siano del tipo complementare descritto nel paragrafo precedente. Si è detto inoltre che, trattandosi di codici binari, la modulazione in fase coincide con quella in ampiezza, dato che si può pensare ai due segnali m_1 e m_2 come sequenze di +1 e -1 che rappresentano l'ampiezza della portante. Lo schema del trasmettitore è semplice e non viene riportato, essendo molto simile a quello della parte alta della fig. 6, con la differenza che, anziché due diverse frequenze, gli oscillatori locali hanno la medesima frequenza e sono sfasati di 90° . L'espressione matematica del segnale trasmesso pertanto è analoga a quella riportata nella (2.2), da completare per la presenza dei due codici:

$$s(t) = m_1(t) \cdot \cos[2\pi f_0(t)] + m_2(t) \cdot \cos[2\pi f_0(t) + \pi/2]. \quad (5.1)$$

Per ricevere correttamente quanto trasmesso si può usare lo stesso sistema già descritto in fig. 3, che va modificato opportunamente: da un lato infatti la struttura è simile perché si devono rivelare due portanti in quadratura, dall'altro si devono combinare le uscite per sfruttare le proprietà dei codici complementari; lo schema risultante è riportato in fig. 9.

Le relazioni già incontrate valgono ancora, con qualche adattamento; per il ramo I:

$$r_2' = r_1' \cdot p' = \left[\frac{m_1(t-t_r)}{\sqrt{2}} \cdot \cos(2\pi f_0 t + \varphi_r) + \frac{m_2(t-t_r)}{\sqrt{2}} \cdot \cos(2\pi f_0 t + \varphi_r + \pi/2) \right] \cdot \cos(2\pi f_0 t) =$$

$$= \frac{m_1(t-t_r)}{2\sqrt{2}} \cdot [\cos(4\pi f_0 t + \varphi_r) + \cos(\varphi_r)] + \frac{m_2(t-t_r)}{2\sqrt{2}} \cdot [\cos(4\pi f_0 t + \varphi_r + \pi/2) + \cos(\varphi_r + \pi/2)] \quad (5.2)$$

$$r_3'(t) = \frac{m_1(t-t_r)}{2\sqrt{2}} \cdot \cos \varphi_r - \frac{m_2(t-t_r)}{2\sqrt{2}} \cdot \sin \varphi_r. \quad (5.3)$$

Invece per il ramo Q, si trova:

$$r_3''(t) = \frac{m_1(t-t_r)}{2\sqrt{2}} \cdot \sin \varphi_r + \frac{m_2(t-t_r)}{2\sqrt{2}} \cdot \cos \varphi_r. \quad (5.4)$$

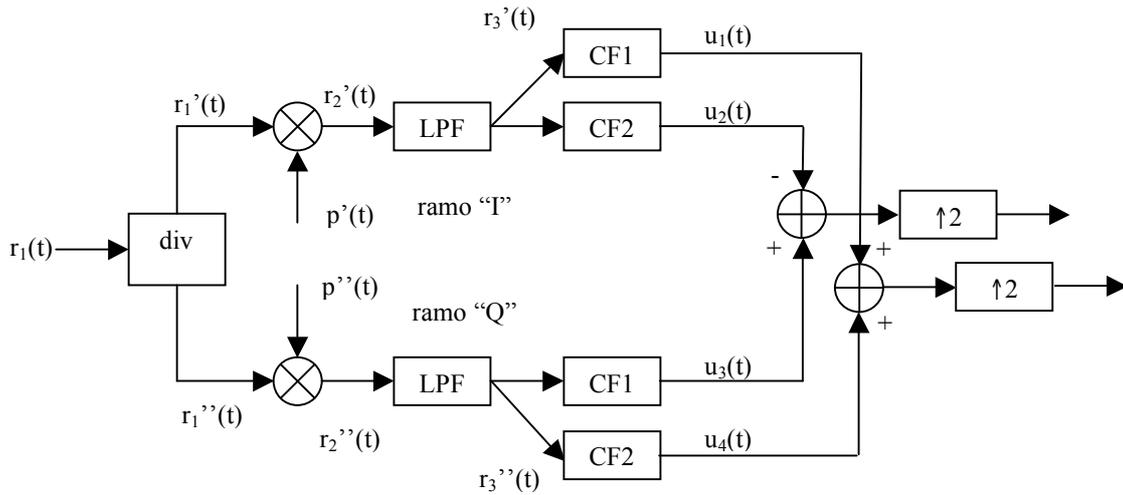


Figura 9. Ricezione di codici complementari con portanti in quadratura (dominio del tempo).

Nei passaggi successivi, al fine di evitare di portare avanti parametri poco significativi, non verranno più riportati il coefficiente di ampiezza $2\sqrt{2}$ e la dipendenza dei codici m_1 e m_2 da $(t-t_r)$. Facendo passare i segnali nei vari filtri a correlazione, si ottengono le uscite seguenti:

$$u_1 = r_3' \circ m_1 = m_1 \circ m_1 \cdot \cos \varphi_r - m_2 \circ m_1 \cdot \sin \varphi_r, \quad (5.5)$$

$$u_2 = r_3' \circ m_2 = m_1 \circ m_2 \cdot \cos \varphi_r - m_2 \circ m_2 \cdot \sin \varphi_r, \quad (5.6)$$

$$u_3 = r_3'' \circ m_1 = m_1 \circ m_1 \cdot \sin \varphi_r + m_2 \circ m_1 \cdot \cos \varphi_r, \quad (5.7)$$

$$u_4 = r_3'' \circ m_2 = m_1 \circ m_2 \cdot \sin \varphi_r + m_2 \circ m_2 \cdot \cos \varphi_r, \quad (5.8)$$

$$u_1 + u_4 = (m_1 \circ m_1 + m_2 \circ m_2) \cdot \cos \varphi_r + (m_1 \circ m_2 - m_2 \circ m_1) \cdot \sin \varphi_r, \quad (5.9)$$

$$u_3 - u_2 = (m_1 \circ m_1 + m_2 \circ m_2) \cdot \sin \varphi_r - (m_1 \circ m_2 - m_2 \circ m_1) \cdot \cos \varphi_r. \quad (5.10)$$

Ricordando le (4.1) e (4.3), si nota che il contributo $(m_1 \circ m_1 + m_2 \circ m_2)$ che compare nella (5.9) o (5.10) è proprio il risultato che, avendo usato codici complementari, porta alla cancellazione dei lobi laterali. Tale cancellazione è perfetta perché i due segnali sono stati trasmessi alla stessa frequenza, tempo e polarizzazione; vi è però il problema del fattore di ampiezza $\cos \varphi_r$. In pratica si ritrova in questo caso lo stesso problema che si era incontrato quando si andava a demodulare un solo codice col sistema omodina, senza sapere la fase della portante ricevuta (relazione (2.10)). Per visualizzare ciò che succede è utile riprendere i diagrammi vettoriali come quelli di fig.4, che nel caso presente si modificano come in fig. 10: in pratica ogni codice risulta affetto da un fattore di ampiezza dovuto al ritardo dell'eco.

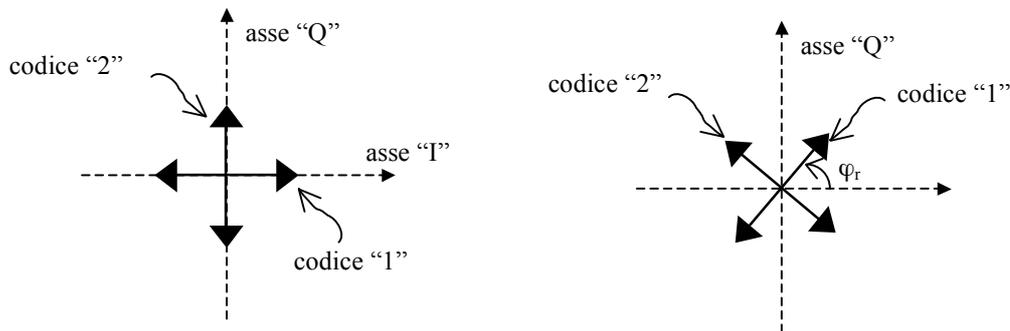


Figura 10. Diagrammi vettoriali dei codici su portanti in quadratura.

In questo caso la cosa migliore che si può tentare per risolvere il problema è elevare al quadrato i risultati e sommare, ottenendo:

$$(u_1 + u_4)^2 + (u_3 - u_2)^2 = (m_1 \circ m_1 + m_2 \circ m_2)^2 + (m_1 \circ m_2 - m_2 \circ m_1)^2. \quad (5.11)$$

A questo punto sono sparite le dipendenze da φ_r ; notare anche che il primo quadrato a secondo membro è applicato ad una somma che già rappresenta la somma delle autocorrelazioni esente da lobi laterali. Il problema è che, oltre al termine dato dalla somma delle autocorrelazioni, compare un termine dato dalla differenza tra le correlazioni incrociate, che non è pari a zero poiché, a differenza della convoluzione, il prodotto di correlazione non gode della proprietà commutativa. La conseguenza è che, anche cercando un codice che minimizzi il problema, si avranno sempre all'uscita dei residui non voluti, dovuti alle correlazioni incrociate.

Per avere un'idea di quanto può succedere in pratica, nel grafico di fig. 11 è riportata una simulazione dell'andamento temporale del termine contenente la somma delle auto correlazioni (ACRS in rosso) e quello delle correlazioni incrociate (CCRD in blu). I due contributi sono rappresentati sovrapposti, ma l'uscita sarebbe unica, data dalla somma delle due, quindi i lobi laterali non sarebbero nulli. Per maggiore precisione, bisognerebbe quadrare, sommare e poi calcolare la radice quadrata, tuttavia quasi ovunque una delle due funzioni è nulla, pertanto si ottiene un analogo risultato sommando i valori assoluti riportati nella figura.

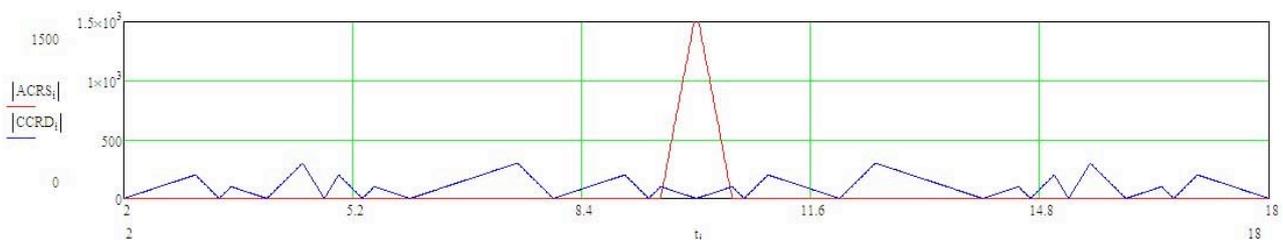


Figura 11. Esempio di uscita del sistema di fig. 9.

I codici usati avevano 16 bit, pertanto era da attendersi un miglioramento nell'ampiezza dell'uscita pari a 32 (unità lineari) ossia 30 dB (guadagno di processo). Un analogo codice a 16 bit, senza cancellazione dei lobi laterali, avrebbe potuto fornire un guadagno di processo di 24 dB ed un rapporto picco-lobi dello stesso valore (se si suppongono lobi laterali unitari). Nel caso di fig. 11 il rapporto picco-lobi è solo di 14.5 dB, pertanto è peggiore di ciò che si sarebbe potuto ottenere con un codice singolo.

Se volessimo tentare di elaborare le informazioni nel dominio della frequenza incontreremmo le stesse problematiche già incontrate nel caso di codice singolo nel paragrafo 3.2. Si dovrebbe correlare l'uscita della FFT una volta col codice 1 e un'altra col codice 2, tuttavia anche ora, in presenza di più echi, la situazione diventerebbe irrisolvibile per la sovrapposizione di φ_r differenti. L'unica speranza di separare gli echi l'uno dall'altro potrebbe essere riposta nella eventualità che essi siano affetti da differenti spostamenti Doppler: se a ciascun eco competesse una velocità differente, si potrebbe sperare di poterne separare gli spettri con opportuno filtraggio. Un procedimento del genere è quanto previsto dal radar di futura generazione di cui in [Bianchi et al., 2008], tuttavia il sistema ivi descritto è molto particolare, dato che usa segnali a banda molto stretta e sfrutta informazioni a priori sulla posizione del bersaglio. In altri casi di interesse geofisico, come il radar glaciologico, le larghezze di banda usate e le velocità tipiche dei bersagli rendono completamente sovrapposti gli spettri degli echi ricevuti.

I problemi incontrati suggeriscono che essi non derivino dal modo in cui viene effettuata la demodulazione, vale a dire che non è il sistema di fig.9 ad essere limitato, ma è la modulazione con portanti in quadratura a non consentire il recupero di tutta l'informazione trasmessa. Per rendersene conto si può concepire un sistema che è la generalizzazione di quello di fig. 9 in cui l'uscita della coppia di mixer data dai due segnali r_2' e r_2'' è pensata come un unico segnale complesso che viene fatto passare attraverso un filtro a correlazione, il quale a sua volta possa correlare l'ingresso con una qualunque combinazione lineare dei due codici, anche a coefficienti complessi. L'uscita di tale filtro dovrebbe avere una forma del genere:

$$r_{3C} = e^{jq_r} [m_1(t - t_r) + j \cdot m_2(t - t_r)] \circ [\alpha \cdot m_1(t) + j \cdot \beta \cdot m_2(t)]. \quad (5.12)$$

In essa il termine esponenziale rappresenta la rotazione di fase della portante, come si presenta all'uscita dei mixer. All'interno della prima parentesi quadra c'è il contributo ricevuto, nella seconda parentesi quadra c'è una combinazione lineare dei due codici, con coefficienti, reali o complessi, da determinare. Svolgendo i calcoli si trova (per semplicità omettiamo l'esponenziale):

$$\alpha \cdot m_1 \circ m_1 + \beta \cdot m_1 \circ m_2 + j \cdot \alpha \cdot m_2 \circ m_1 + j \cdot \beta \cdot m_2 \circ m_2,$$

che mostra chiaramente che l'unica scelta possibile per annullare il contributo delle correlazioni incrociate è $\alpha=\beta=0$, che però annulla anche le auto correlazioni.

Generalizzando ulteriormente, si potrebbe mostrare che perfino operando direttamente ad IF, si otterrebbe lo stesso risultato, vale a dire che operando su $r_1(t)$ ed effettuando la correlazione (ovviamente con una versione del codice su portante) si otterrebbe come risultato una somma di segnali con spettro non più contiguo alla frequenza zero ma a quella della portante a IF. Tra questi segnali ci sarebbe la portante a IF (peraltro facilmente eliminabile), oltre al risultato della correlazione tra i codici 1-1, 2-2, 1-2 e 2-1, di nuovo non separabili. La formulazione matematica viene tralasciata per non appesantire inutilmente il discorso.

L'analisi effettuata in questo paragrafo è servita solo a mettere in evidenza ciò che era stato anticipato alla fine del paragrafo precedente. Il fatto che in uscita si ritrovi la sovrapposizione dei segnali trasmessi sulle due portanti deriva da una proprietà intrinseca del sistema di trasmissione, che non sfrutta due canali in qualche modo separabili, a meno che non si disponga dell'informazione della fase della portante. L'unico modo per realizzare la trasmissione contemporanea dei due codici è sfruttare un doppio canale, come nel sistema che viene analizzato nel prossimo paragrafo.

6. Sfruttamento di due canali con polarizzazione circolare opposta

Si è accennato in precedenza alla possibilità di usare la polarizzazione incrociata per trasmettere i due codici complementari contemporaneamente. Si è detto anche che, qualora i piani di polarizzazione ruotino (anche rimanendo ortogonali), si avrebbero problemi di interferenza tra i due canali, analoghi a quelli incontrati nel caso della modulazione con portanti in quadratura. Un metodo che eviterebbe i problemi legati alla citata rotazione dei piani di polarizzazione dovrebbe fare uso della polarizzazione circolare. È noto che per ottenere una polarizzazione del genere è necessario trasmettere con due antenne lineari due portanti opportunamente sfasate. La trasmissione contemporanea dei due codici si avrebbe usando polarizzazioni circolari con versi di rotazione opposti, i quali non si influenzano, come si vedrà tra breve. In fig. 12 è riportato lo schema di principio del trasmettitore.

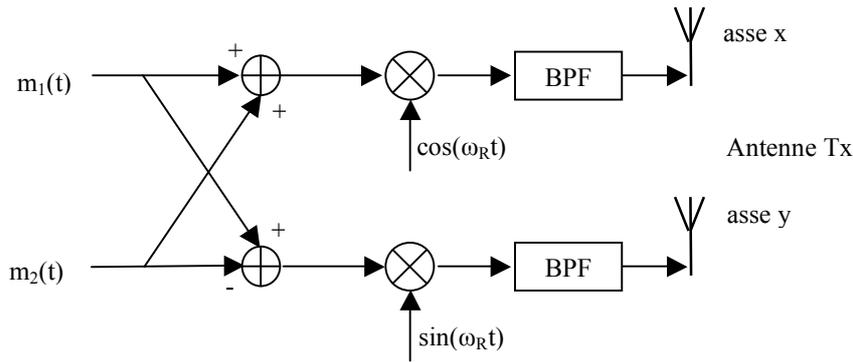


Figura 12. Trasmissione in polarizzazione circolare.

Il vettore risultante di uno dei campi (per esempio quello elettrico) si propaga lungo l'asse z , perpendicolare al piano formato dalle due antenne, orientate secondo l'asse x e y rispettivamente. Fissato un punto dell'asse z , per esempio per $z=0$, l'andamento nel tempo si può scrivere come:

$$\vec{E} = (m_1 + m_2) \cos(\omega_R t) \vec{i} + (m_1 - m_2) \sin(\omega_R t) \vec{j}, \quad (6.1)$$

ove i versori degli assi x e y sono indicati con \vec{i} e \vec{j} in grassetto oppure con una freccetta sovrapposta. Il vettore risultante si può immaginare così decomposto:

$$\begin{aligned} \vec{E} &= \vec{E}_{ccw} + \vec{E}_{cw} \\ \vec{E}_{ccw} &= m_1 \cdot [\cos(\omega_R t) \vec{i} + \sin(\omega_R t) \vec{j}], \\ \vec{E}_{cw} &= m_2 \cdot [\cos(\omega_R t) \vec{i} - \sin(\omega_R t) \vec{j}] \end{aligned} \quad (6.2)$$

con le due componenti \vec{E}_{ccw} e \vec{E}_{cw} che ruotano in versi opposti, a causa delle particolari relazioni di fase delle funzioni seno e coseno che compaiono nelle loro espressioni. I due codici sembrano sovrapporsi, così almeno appare essere la proiezione sui due assi del campo elettrico risultante. In realtà, a differenza del caso delle portanti in quadratura, in questo caso non è necessario disporre di un riferimento di fase. Ovviamente, per separare correttamente i due canali bisogna fare attenzione alla struttura del ricevitore, riportata in fig. 13 (i nomi dei segnali sono stati dati in analogia con le figure precedenti).

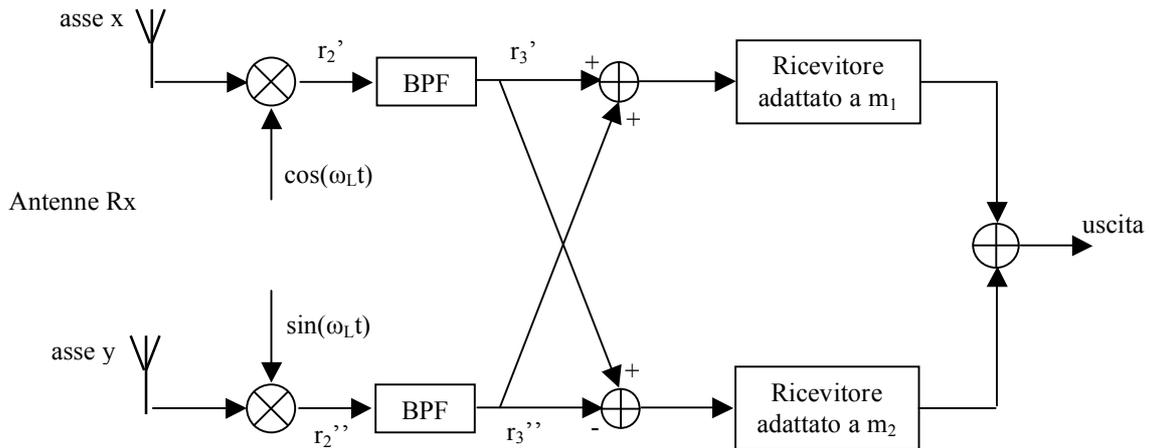


Figura 13. Ricezione in polarizzazione circolare.

Come sottolineato a proposito del sistema di fig. 6, è necessario considerare ciò che accade anche a media frequenza (e quindi in presenza di un oscillatore locale con pulsazione ω_L). All'uscita dei mixer che convertono a media frequenza si hanno:

$$\begin{aligned} r_2'(t) &= [m_1(t - t_r) + m_2(t - t_r)] \cos(\omega_R t + \varphi_r) \cdot \cos(\omega_L t) \\ r_2''(t) &= [m_1(t - t_r) - m_2(t - t_r)] \sin(\omega_R t + \varphi_r) \cdot \sin(\omega_L t) \end{aligned} \quad (6.3)$$

come fatto in precedenza, si è indicato con t_r il ritardo dell'eco e con φ_r lo sfasamento sulla portante dovuto allo stesso ritardo. Trascurando i fattori di ampiezza dovuti all'applicazione delle formule di Werner e i termini $(t-t_r)$ della dipendenza temporale di m_1 e m_2 , all'uscita dei filtri passa banda si trova:

$$\begin{aligned} r_3'(t) &= (m_1 + m_2) \cos[(\omega_R - \omega_L)t + \varphi_r] \\ r_3''(t) &= (m_1 - m_2) \cos[(\omega_R - \omega_L)t + \varphi_r] \end{aligned} \quad (6.4)$$

ed infine, all'uscita dei sommatore:

$$\begin{aligned} r_3'(t) + r_3''(t) &= m_1 \cdot \cos[(\omega_R - \omega_L)t + \varphi_r] \\ r_3'(t) - r_3''(t) &= m_2 \cdot \cos[(\omega_R - \omega_L)t + \varphi_r] \end{aligned} \quad (6.5)$$

In definitiva, ciò che si presenta a ciascun ingresso dei blocchi "ricevitore adattato" dipende solo da uno dei due codici. È vero che si ritrova uno sfasamento φ_r sulla portante a media frequenza, ma esso è del tutto analogo a quello che si era incontrato nel sistema a codice singolo, pertanto, se in sede di demodulazione in banda base si adotta una delle soluzioni già prospettate (figg. 3 o 5a), è possibile estrarre i due codici al massimo delle loro ampiezze. L'uscita è il risultato della somma dei due codici e quindi presenta un picco centrale senza lobi laterali.

Anche se elegante, la soluzione esposta in questo paragrafo ha anch'essa il suo limite, in quei casi in cui il mezzo non isotropo determini diverse caratteristiche di propagazione, come la ionosfera. In questi casi i segnali che viaggiano con le due polarizzazioni possono arrivare al ricevitore con ritardi differenti, con l'evidente degradazione dei risultati. (Le differenti modalità di propagazione delle polarizzazioni con verso opposto sono alla base degli studi delle proprietà della ionosfera; si veda ad esempio [Sciacca e al., 2007].) La ricerca di una trasmissione contemporanea dei due codici infatti era partita proprio dalla necessità di evitare tempi di ritardo differenti, possibili nel caso di trasmissione in sequenza. Per fortuna gli echi della ionosfera possono considerarsi quasi fermi perciò il metodo a "divisione di tempo" può applicarsi con

successo. Il metodo che usa la polarizzazione circolare potrebbe invece essere usato nei casi di movimento rapido in mezzi isotropi.

7. Considerazioni sulle ampiezze dei segnali ricevuti

Un tema sul quale si è sorvolato, introducendo i vari sistemi per trattare i codici, riguarda le ampiezze dei segnali. Si è ipotizzato che il valore di ampiezza non sia importante, purché sufficiente a far sì che il ricevitore rilevi la presenza di un oggetto. Per evitare che i livelli in ricezione si abbassino troppo, è compito del progettista del sistema assicurare un livello di trasmissione adeguato, antenne efficienti ed un ricevitore che amplifichi i segnali captati quanto basta perché possano essere elaborati. Per questi motivi non ci si è preoccupati molto di porre un fattore di ampiezza davanti ai vari termini che comparivano nelle varie formule.

Ciò nonostante, qualche coefficiente talvolta è comparso ed è stato mantenuto: si tratta del divisore per due che compare in seguito alla demodulazione, risultato dell'applicazione delle formule di Werner, e del divisore per radice di due che tiene conto della suddivisione del segnale di ingresso al ricevitore a due canali. Questi coefficienti sono stati lasciati indicati perché ad essi corrisponde effettivamente una perdita di energia, che è utile considerare per confrontare tra loro sistemi diversi.

Tale confronto porta ad un risultato a prima vista deludente: il sistema a canale singolo di fig. 2 presenta in uscita un'ampiezza proporzionale ad $1/2$ del picco determinato dalla funzione $a(t)$ (relazione 2.10), mentre nel sistema a doppio canale di fig. 3 l'uscita è proporzionale a $1/\sqrt{8}$ (relazione 3.6, dopo estrazione di radice). In pratica sembra che si abbia una perdita, dovuta al separatore dei due canali, che non viene più recuperata. In realtà in questo ragionamento non si è tenuto conto del fatto che l'uscita del sistema a canale singolo dipende da $\cos\varphi$, quindi fluttua e non assume mai il valore massimo $1/2$ prima detto.

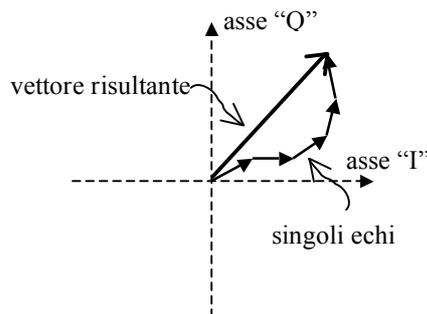


Figura 14. Composizione dei vettori relativi ad echi con fase diversa.

Volendo fare un confronto tra le ampiezze dei segnali ricevuti dai due sistemi bisognerebbe calcolare una media tra tutti i valori dipendenti da $\cos\varphi$. Per fare ciò si deve elaborare un modello, sia pure semplificato, di ciò che succede al segnale nel sul viaggio dal trasmettitore al riflettore e ritorno. All'uopo si deve tener presente che non solo il mezzo può presentare disomogeneità che possono far cambiare la lunghezza del cammino percorso dall'onda, ma lo stesso riflettore è sempre costituito da svariati elementi riflettenti, ciascuno posto ad una distanza leggermente differente dagli altri; infine bisognerebbe tener conto anche delle differenti proprietà riflettenti di tali elementi, che portano a differenti quantità di energia riflessa.

Ciò che si presenta all'ingresso del ricevitore, quindi, è un segnale risultato dalla composizione di vari vettori, ciascuno corrispondente ad un riflettore elementare (vedi fig. 14). La somma vettoriale dà luogo ad un segnale ricevuto che avrà un'ampiezza ed una fase fluttuanti, più o meno lentamente (difficilmente il numero di riflettori elementari che concorrono a formare l'eco è così elevato da portare ad una somma nulla, ciò è anche dovuto alla dimensione della cella di risoluzione in distanza che evidentemente discrimina gli echi che si sovrappongono da quelli che cadono altrove).

Ora, il nostro sistema ricevente non vede i singoli vettori componenti e non può fare altro che tentare la rivelazione del vettore risultante. Limitandoci per ora agli effetti della variazione di fase (supponendo quindi che l'ampiezza non vari), ci si propone di stimare una media dei possibili valori ricevuti dal sistema a canale singolo, che fornisce in uscita un'ampiezza proporzionale a $\cos\varphi$. Essendo interessati al modulo del segnale

all'uscita del ricevitore, anziché far variare φ tra $-\pi$ e π è sufficiente farlo variare tra $-\pi/2$ e $\pi/2$. Si trova (k è una costante indipendente dall'angolo):

$$\bar{u} = k \cdot \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} \cos \varphi \cdot d\varphi = \frac{k}{\pi} \cdot [\sin \varphi]_{-\pi/2}^{\pi/2} = \frac{2k}{\pi}. \quad (7.1)$$

Cioè l'ampiezza è $2/\pi$ volte quella che si sarebbe avuta se $\cos \varphi$ fosse stato sempre pari ad uno; a questo fattore va applicato il fattore $1/2$ di cui si parlava all'inizio. In definitiva, in media, il sistema a singolo canale ha un'uscita proporzionale a $1/\pi$, mentre per il sistema a doppio canale è proporzionale a $1/\sqrt{8}$; il rapporto tra i due è di circa 1.11 a favore del sistema a doppio canale, che corrisponde a circa 1 dB.

Il fattore di merito appena ricavato può sembrare tale da non giustificare la costruzione di un sistema a doppio canale; vi è da considerare però che il valore del segnale ricevuto è molto più stabile rispetto al sistema a canale singolo, che può fornire ad ogni ricezione un valore variabile da zero a $1/2$. Inoltre alle variazioni dell'ampiezza in uscita valutate sopra andrebbero aggiunte quelle dovute alla variazione di ampiezza degli echi (causate dalla differente combinazione dei vettori componenti della fig. 14 e alle quali nessuno dei due sistemi descritti è in grado di porre rimedio). Per ovviare alle fluttuazioni del segnale si devono adottare metodi di integrazione su più impulsi; in tale contesto, avere a disposizione impulsi da integrare che sono più stabili diminuisce il tempo di integrazione rispetto al sistema a canale singolo.

8. Conclusioni

In questo lavoro sono stati esaminati i vari modi in cui è possibile trasmettere due codici di fase complementari, il cui uso consente, almeno in linea di principio, di annullare il problema dei lobi laterali del quale sono affetti i codici di fase ordinari. Operando una schematizzazione, i risultati possono essere riassunti come segue.

Qualora i bersagli illuminati dal radar siano fermi, o si muovano molto lentamente, come nel caso dei sondaggi ionosferici, il metodo migliore è senz'altro la trasmissione consecutiva dei due codici, la loro memorizzazione, e la loro somma dopo la ricezione di entrambi.

Nel caso in cui i bersagli dovessero muoversi relativamente veloci, rispetto alle dimensioni della cella di risoluzione in distanza, la non perfetta sincronizzazione degli echi potrebbe far ricomparire i lobi laterali. In questi casi si può scegliere una delle due soluzioni: se la degradazione è accettabile, è preferibile tenersela, se non addirittura optare per l'uso di un codice solo (senza cancellazione dei lobi laterali), che almeno comporta un hardware più semplice. Se la degradazione fosse eccessiva, purché in presenza di mezzi isotropi, si può prendere in considerazione il metodo della polarizzazione circolare, purché il maggior costo dell'hardware sia accettabile a fronte dei miglioramenti delle prestazioni.

Bibliografia

- Baskaradas J.A., Bianchi C., Sciacca U., Tutone G., Zuccheretti E. (2002). *The new AIS INGV digital ionosonde design report*. Rapporti Tecnici INGV, n°12.
- Bianchi C., Sciacca U., Arokiasami, J.B., Zuccheretti, E., Tutone, G. (2008). *Rilevamento delle sorgenti doppler nella ionosfera tramite radar HF*, Rapporti Tecnici INGV, n°62.
- Cabrera, M.A., Zuccheretti E., Ezquer R.G., Sciacca U., et al. (2010). *Some considerations for different time-domain signal processing of pulse compression radar*. Annals of Geophysics, vol.53, n°5-6.
- Sciacca U., Bianchi C., Zuccheretti E. (2007). *L'uso della polarizzazione dell'eco nei sondaggi ionosferici*. Quaderni di Geofisica, n°50.
- Skolnik, M. (1990). *Radar Handbook*. McGraw-Hill, New York. 2nd edition.
- Zirizzotti A., Baskaradas J.A., Bianchi C., Sciacca U., Tabacco I.E., Zuccheretti E. (2008). *Glacio Radar system and results*. 2008 IEEE Radar Conference paper.

Coordinamento editoriale e impaginazione

Centro Editoriale Nazionale | INGV

Progetto grafico e redazionale

Daniela Riposati | Laboratorio Grafica e Immagini | INGV

© 2011 INGV Istituto Nazionale di Geofisica e Vulcanologia

Via di Vigna Murata, 605

00143 Roma

Tel. +39 06518601 Fax +39 065041181

<http://www.ingv.it>



Istituto Nazionale di Geofisica e Vulcanologia