

FORMAREA DIAGramei DE RADIAȚIE A ANTENELOR

1.1 Fundamente ale teoriei șirurilor de antene

Un șir de antene este un ansamblu de antene identice, identic orientate și cu distribuții de curent similare (prin similaritate se înțelege că distribuțiile de curent diferă între ele numai prin coeficienți complecși multiplicativi). Se definește *antena de referință* a șirului ca fiind o antenă identică cu antenele din șir, identic orientată cu acestea și având distribuție de curent similară cu particularitatea că *un punct al ei reprezintă originea sistemului de coordonate*.

Potrivit definiției de mai sus și notațiilor din figura 1.1 o antenă m oarecare din șir se obține din antena de referință prin translație cu un vector \mathbf{r}_m , iar distribuția ei de curent – prin multiplicarea distribuției de curent a antenei de referință cu un coeficient complex w_m . Analitic:

$$\mathbf{r}'_m = \mathbf{r} + \mathbf{r}_m, \quad J(\mathbf{r}'_m) = w_m J_{ref}(\mathbf{r}') \quad (1.1)$$

Dacă antenele din șir sunt distribuite după o dreaptă șirul se numește liniar, iar dacă sunt distribuite pe circumferința unui cerc șirul se numește circular. Când antenele din șir ocupă o suprafață într-un plan șirul poartă denumirea formeii suprafeței, respectiv: hexagonal, rectangular, triunghiular etc. Există și șiruri ce ocupă suprafețe în spațiu sau volume compacte, dar acestea nu au aplicații în rețelele de comunicații mobile.

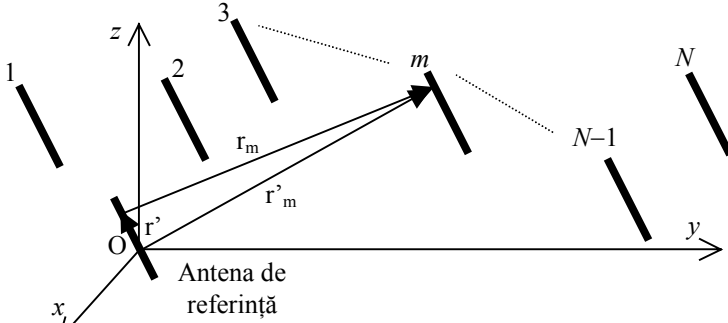


Figura nr. 1.1 – Arhitectura unui șir de antene

Dacă $|E_{ref}(\theta, \phi)|$ este diagrama de radiație normalizată a antenei de referință (identică, de fapt, cu diagrama de radiație a oricărei antene din șir), atunci diagrama de radiație a șirului de antene poate fi scrisă ca:

$$|E_{șir}(\theta, \phi)| = |E_{ref}(\theta, \phi)| \cdot |f(\theta, \phi)| \quad (1.2)$$

unde $f(\theta, \phi)$ se numește *factor de șir* și este o funcție ce depinde numai de poziția relativă a antenelor din șir și de coeficienții complecși w_m . În context $|E_{ref}(\theta, \phi)|$ se numește *factor de element*. Se observă că diagrama totală de radiație a șirului se obține prin multiplicarea diagramei de radiație a antenelor din șir cu modulul factorului de șir.

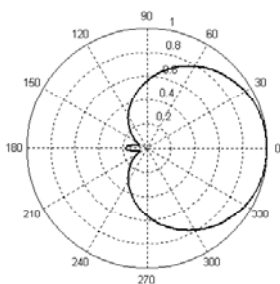
Se poate demonstra că:

$$f(\theta, \phi) = \sum_{m=1}^N w_m e^{jk \mathbf{r}_v \cdot \mathbf{r}_m} \quad (1.3)$$

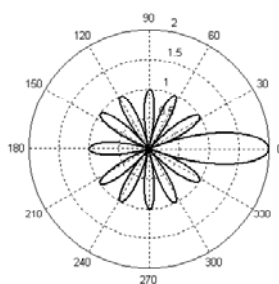
unde $k_0 = 2\pi / \lambda_0$ este constanta de propagare (numărul de undă) pentru spațiul liber, iar \mathbf{r}_v este versorul corespunzător direcției de propagare considerate.

De regulă, antenele din șir (deci, și antena de referință) au o diagramă de radiație cu un lob principal de deschidere mare și cu lobi secundari foarte mici sau inexistenți. Factorul de șir are o formă mult mai complicată, cu un lob principal foarte îngust și cu numeroși lobi secundari.

El are rolul de a „modula” distribuția spațială a radiației antenelor astfel ca în ansamblu șirul să-și concentreze radiația într-un unghi spațial îngust și orientat în direcția dorită. De reținut că, potrivit relației (1.3), aceasta se obține prin poziționarea relativă adecvată (vectorii \mathbf{r}_m) a antenelor din șir și alimentarea corespunzătoare (coeficienții w_m) a lor. În figura 1.2 sunt prezentate diagrame tipice de radiație pentru antenele din structura unui șir și pentru factorul de șir.



antena de referință



factorul de șir

Figura nr. 1.2 – *Diagrame de radiație tipice*

Dacă antenele din șir sunt echidistant așezate, au același modul al distribuțiilor de curent, iar faza acestora variază cu aceeași cantitate de la o antenă la alta șirul se numește *uniform*. Pentru analize teoretice și aplicații practice mai simplu de implementat de foarte mare importanță sunt *șirurile liniare uniforme* și *șirurile circulare uniforme*.

Șiruri liniare

Fie Oz axa unui sistem cartezian de coordonate după care sunt distribuite echidistant antenele unui șir liniar. Notăm cu Δ distanța dintre două antene alăturate (denumită și *constanta de distanță* a șirului) și cu δ – cantitatea cu care variază faza de la o antenă la alta (denumită și *constanta de fază* a șirului). Nu impunem condiții asupra modulelor distribuțiilor de curent, deci șirul nu este uniform. Considerăm că șirul are N antene (denumite și elemente ale șirului) și că primul element din șir este antenă de referință. În aceste condiții:

$$\mathbf{r}_m = (m-1) \Delta \mathbf{z}_v \quad \text{și} \quad \delta_m = (m-1) \delta, \quad m = 1 \dots N \quad (1.4)$$

și, ca urmare:

$$f(\theta, \phi) = \sum_{m=1}^N w_m e^{jk_0(m-1)\Delta \cos \theta} = \sum_{m=1}^N |w_m| e^{j(m-1)(\delta + k_0\Delta \cos \theta)} \quad (1.5)$$

Se observă că dacă se dorește ca lobul principal al diagramei de radiație a șirului să fie orientat în direcția θ_0 , de exemplu, atunci este necesar ca șirul să aibă o constantă de fază:

$$\delta = -k_0 \Delta \cos \theta_0 \quad (1.6)$$

Cu notațiile:

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_N \end{bmatrix} \quad \mathbf{v} = \begin{bmatrix} 1 \\ e^{jk_0\Delta \cos \theta} \\ e^{j2k_0\Delta \cos \theta} \\ \vdots \\ e^{j(N-1)k_0\Delta \cos \theta} \end{bmatrix} \quad (1.7)$$

expresia (1.5) a factorului de șir se poate scrie compact sub forma:

$$f(\theta, \phi) = \mathbf{w}^T \cdot \mathbf{v} \quad (1.8)$$

unde T înseamnă transpunere.

Șiruri circulare

Dacă cele N elemente ale unui șir sunt distribuite echidistant pe circumferința unui cerc de rază R în planul xOy al unui sistem cartezian având centrul în centrul sistemului de coordonate, dacă elementul cu numărul 1 este așezat pe axa Ox și faza distribuției lui de curent este considerată fază de referință și dacă antena de referință se află în centrul cercului, atunci:

$$\mathbf{r}_m = R \cos \phi_m \mathbf{x}_v + R \sin \phi_m \mathbf{y}_v \quad \phi_m = (m-1) \frac{2\pi}{N} \quad m = 1 \dots N \quad (1.9)$$

și, deci:

$$\mathbf{r}_v \cdot \mathbf{r}_m = R \sin \theta \cos(\phi - \phi_m) \quad (1.10)$$

astfel că:

$$f(\theta, \phi) = \sum_{m=1}^N w_m e^{jk_0 R \sin \theta \cos(\phi - \phi_m)} = \sum_{m=1}^N |w_m| e^{j[\delta_m + k_0 R \sin \theta \cos(\phi - \phi_m)]} \quad (1.11)$$

Pentru ca lobul principal al diagramei de radiație a șirului să fie orientat în direcția (θ_0, ϕ_0) este necesar ca:

$$\delta_m = -k_0 R \sin \theta_0 \cos(\phi_0 - \phi_m) \quad (1.12)$$

1.2 Formarea lobului (beamforming)

Semnificația inițială a tehnicii de formare a diagramei de radiație a unei antene (pe scurt: *formare a lobului*) consta în orientarea lobului principal al diagramei de radiație a unei antene cu reflector parabolic în direcția permanent schimbătoare a unui satelit de comunicații. Odată cu aplicarea acestei tehnici șirurilor de antene cu control electronic al radiației s-a îmbogățit și semnificația ei, ea incluzând actualmente și crearea unor nului ale diagramei de radiație în direcțiile din care sosesc semnalele de interferență și, eventual, menținerea nivelului lobilor secundari sub un prag prestabilit.

Formarea analogică a lobului

Primele implementări ale tehnicii de formare a lobului sau făcut cu blocuri analogice, cele mai semnificative fiind cele ce utilizează matrici Butler pentru controlul fazei distribuțiilor de curent ale antenelor din șir. O matrice Butler este un ansamblu de cuploare hibride și circuite de defazare care furnizează la diverse ieșiri semnale cu relație de fază controlată, dar fixă. Prin interconectarea mai multor matrici Butler se obțin simultan la diverse porturi de ieșire diagrame de radiație independente, eventual ortogonale între ele, ale aceluiași șir de antene.

De exemplu, folosind patru matrici Butler interconectate ca în figura 1.3 se obțin patru orientări distincte ale lobului principal la cele patru porturi de ieșire. În Tabelul 1.1 sunt prezentate fazele coeficienților complecși w_m , considerând faza primului element ca referință.

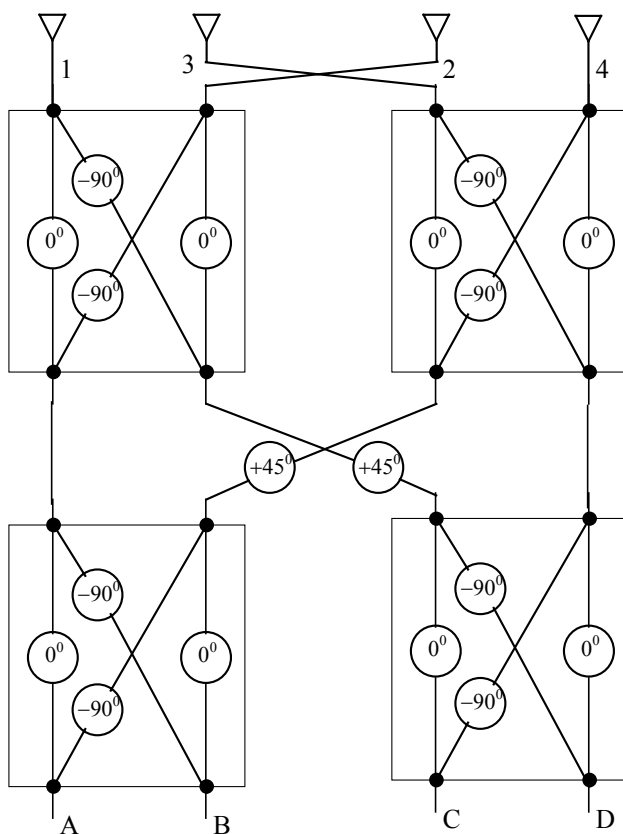


Figura nr. 1.3 - Circuit de formare analogică a lobului cu matrici Butler

Tabelul nr. 1.1 – Fazele coeficienților de ponderare w_m la ieșirile circuitului de formare analogică a lobului cu matrici Butler

Portul de ieșire	Defazajul introdus pe calea semnalului furnizat de antena nr.:				Faza relativă a semnalului furnizat de antena nr.:				Constanta de fază a șirului
	1	2	3	4	1	2	3	4	
A	0	-45^0	-90^0	-135^0	0	-45^0	-90^0	-135^0	-45^0
B	-90^0	$+45^0$	-180^0	-45^0	0	$+135^0$	$+270^0$	$+405^0$	$+135^0$
C	-45^0	-135^0	$+90^0$	-45^0	0	-135^0	-270^0	-405^0	-135^0
D	-135^0	-90^0	-45^0	0	0	$+45^0$	$+90^0$	$+135^0$	$+45^0$

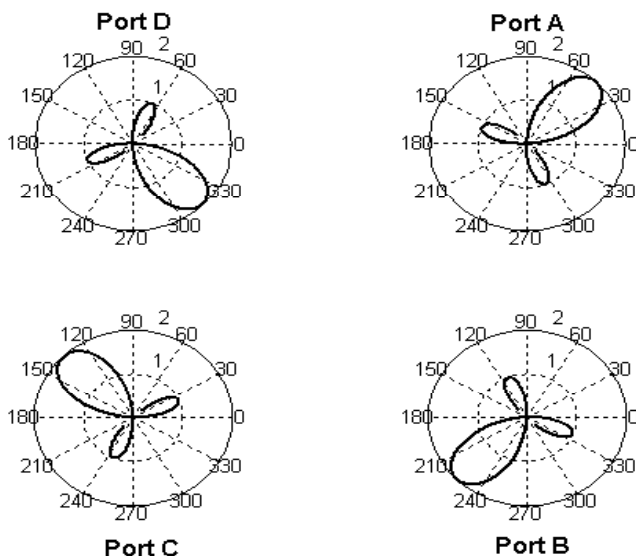


Figura nr. 1.4 – *Diagramele de radiație ale șirului la porturile de ieșire ale formatorului analogic cu matrici Butler*

Dacă distanța dintre elementele șirului este $\lambda_0/2$, atunci diagramele de radiație ale șirului la cele patru porturi au forma prezentată în figura 1.4. De remarcat că cele patru diagrame de radiație ale șirului de antene sunt create simultan, fiecare la unul din porturile de ieșire ale circuitului de formare a lobului.

Configurațiile de circuit cu matrici Butler asigură orientări fixe, prestabilite ale lobului principal. Dacă se dorește ca direcția principală de radiație să se modifice prin comandă electronică trebuie introduse circuite de defazare individuale pentru fiecare antenă din șir și, apoi, să se sumeze semnalele obținute (figura 1.5). Ansamblul alcătuit din circuitele de defazare și sumator formează circuitul propriu-zis de formare a lobului. Deoarece acesta modifică numai faza semnalelor furnizate de antene șirul căruia i se aplică un astfel de control se numește *șir fazat*. Când se dorește obținerea simultană a mai multor direcții ale lobului principal pornind de la același șir de antene trebuie utilizate în paralel mai multe circuite de formare a lobului. Datorită numărului mare de componente ce intră în structura acestuia, raportul semnal/zgomot la ieșire se degradează puternic, acesta

constituind unul din dezavantajele majore ale șirurilor fazate, dezavantaj ce este comun, de fapt, tuturor circuitelor analogice de formare a lobului.

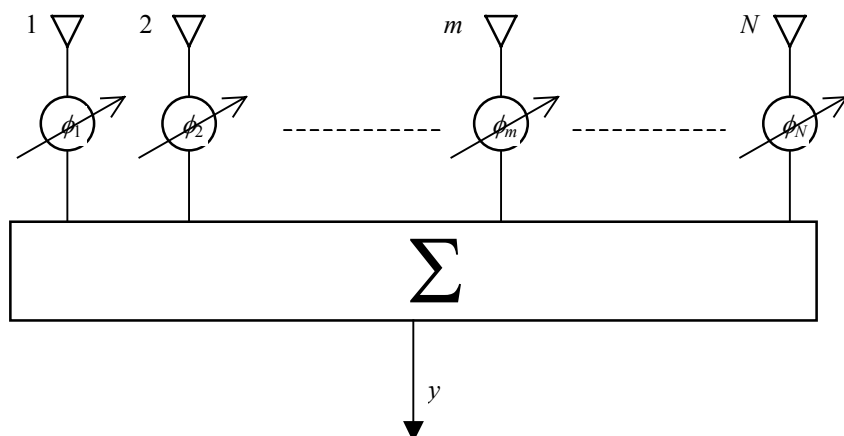


Figura nr. 1.5 – Șir fazat

Formarea digitală a lobului

Utilizarea unui control digital al diagramei de radiație a unui șir de antene presupune reprezentarea numerică a semnalelor furnizate de antenele șirului, semnale ce conțin informația despre distribuția spațială a undelor electromagnetice incidente. În acest scop semnalele individuale sunt eșantionate, iar eșantioanele sunt aplicate unor blocuri de conversie analog/digitală. Dacă frecvența semnalelor recepționate este prea mare pentru operațiunile necesare de eșantionare și conversie A/D, se poate micșora frecvența de lucru prin heterodinare (funcționare în frecvență intermediară).

Reprezentarea digitală a informației aduce cu sine toate avantajele prelucrării numerice, adică: prelucrări rapide și complexe, nealterarea raportului semnal/zgomot, generarea simultană și independentă a mai multor direcții pentru lobul principal, modificarea simplă a algoritmului de prelucrare etc. Formarea digitală a lobului permite introducerea în mod real a controlului adaptiv, lobul principal putând fi orientat, teoretic, în *orice* direcție din spațiu și, în plus, fiind posibilă și crearea de nului de diagramei de radiație în direcțiile din care sosesc semnale de interferență.

Formarea adaptivă a lobului

Un bloc de formare adaptivă a lobului modifică diagrama de radiație a șirului de antene în funcție de configurația mediului electromagnetic în vederea satisfacerii unei funcții obiectiv. Astfel, este posibilă separarea semnalului util de semnalele de interferență din aceeași bandă de frecvență, dar care sosesc din alte direcții decât semnalul util.

Istoric, o primă formă de adaptivitate în formarea lobului s-a implementat prin introducerea anulatorului de lob lateral în sistemele radar și sonare însă conceptul modern de adaptivitate așa cum se folosește și astăzi a fost dezvoltat în anii '60 în mod independent de Appelbaum și, respectiv, Widrow și Hoff. În primul caz criteriul de optimizare l-a constituit maximizarea raportului semnal/zgomot al semnalului de ieșire, iar în al doilea caz – minimizarea erorii pătratice medii. În condiții de staționaritate a mediului electromagnetic soluțiile ambilor algoritmi converg către soluția optimă Wiener. Ulterior a fost dezvoltat criteriul asemănării maxime (maximum likelihood) sau al răspunsului nedistorionat de varianță minimă. Acest algoritm a fost primul care a permis separarea semnalului util de un semnal de interferență a cărui direcție de sosire este la numai o fracțiune din deschiderea lobului principal față de direcția de sosire a semnalului util. În sfârșit, prin dezvoltarea algoritmului bazat pe inversarea matricii de eșantioane se poate calcula direct vectorul ponderilor și se elimină unele probleme de convergență ale algoritmilor anteriori.

Exemplu

Scopul unui algoritm de formare digitală a lobului este acela de a determina valorile optime ale ponderilor complexe cu care trebuie multiplicat semnalele furnizate de antenele individuale pentru ca în semnalul obținut prin sumarea acestora semnalul util să aibă nivel maxim posibil, iar semnalele de interferență – nivel minim posibil. Ilustrăm conceptul de formare digitală adaptivă a lobului cu ajutorul unui șir de antene cu două elemente (figura 1.6) așezate la distanța de $\lambda_0/2$ unul de altul și care recepționează un semnal util din direcția perpendiculară pe axa șirului și un semnal de interferență pe aceeași frecvență, dar dintr-o direcție ce face un unghi de 30° cu direcția de sosire a semnalului util.

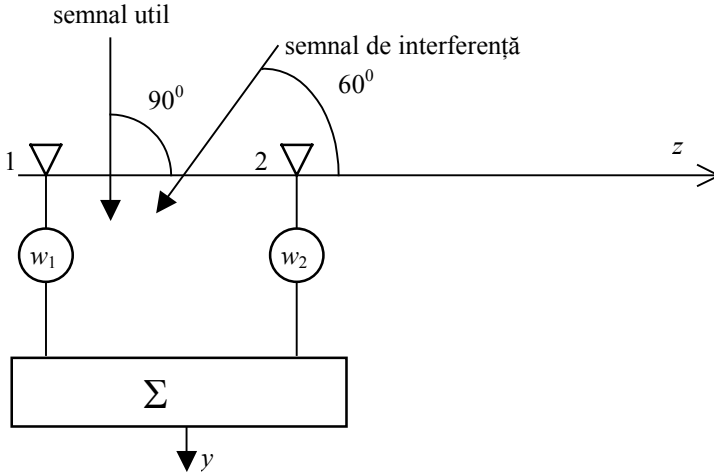


Figura nr. 1.6 – Șir cu două elemente controlat digital

Avem:

$$k_0 \Delta = \frac{2\pi}{\lambda_0} \frac{\lambda_0}{2} = \pi \quad (1.13)$$

Pentru recepția semnalului util $\theta = \theta_u = \pi/2$ și $k_0 \Delta \cos \theta_u = 0$, astfel că, potrivit relației (1.7):

$$\mathbf{v} = \mathbf{v}_u = \begin{bmatrix} 1 \\ e^{j0} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (1.14)$$

iar pentru recepția semnalului de interferență $\theta = \theta_p = \pi/3$ și, $k_0 \Delta \cos \theta_p = \pi/2$, astfel că:

$$\mathbf{v} = \mathbf{v}_p = \begin{bmatrix} 1 \\ e^{j\pi/2} \end{bmatrix} = \begin{bmatrix} 1 \\ j \end{bmatrix} \quad (1.15)$$

Dacă se lucrează cu valori normalizate pentru ponderi ($|w_1| + |w_2| = 1$), atunci modulul factorului de șir (relația 1.3) are valoarea minimă egală cu 0 și valoarea maximă egală cu 1. Pentru satisfacerea criteriului de optimalitate este deci necesar ca:

$$\begin{cases} \mathbf{w}^T \cdot \mathbf{v}_u = [w_1 & w_2] \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = w_1 + w_2 = 1 \\ \mathbf{w}^T \cdot \mathbf{v}_p = [w_1 & w_2] \cdot \begin{bmatrix} 1 \\ j \end{bmatrix} = w_1 + jw_2 = 0 \end{cases} \quad (1.16)$$

sau, echivalent:

$$\begin{cases} \operatorname{Re}(w_1) + \operatorname{Re}(w_2) = 1 & \operatorname{Re}(w_1) - \operatorname{Im}(w_2) = 0 \\ \operatorname{Im}(w_1) + \operatorname{Re}(w_2) = 0 & \operatorname{Im}(w_1) + \operatorname{Im}(w_2) = 0 \end{cases} \quad (1.17)$$

de unde rezultă:

$$w_1 = \frac{1}{2} - j\frac{1}{2} \quad w_2 = \frac{1}{2} + j\frac{1}{2} \quad (1.18)$$

Pentru aceste valori optime ale ponderilor expresia (1.3) a factorului de șir devine:

$$\begin{aligned} f(\theta, \phi) &= \sum_{m=1}^2 w_m e^{j(m-1)k_0 \Delta \cos \theta} = \\ &= w_1 + w_2 e^{j\pi \cos \theta} = \left(\frac{1}{2} - j\frac{1}{2} \right) + \left(\frac{1}{2} + j\frac{1}{2} \right) e^{j\pi \cos \theta} \end{aligned} \quad (1.19)$$

cu:

$$|f(\theta, \phi)| = \sqrt{2} \left| \sin \left(\frac{\pi}{4} - \frac{\pi \cos \theta}{2} \right) \right| \quad (1.20)$$

Din reprezentarea grafică a expresiei de mai sus (figura 1.7) se observă că diagrama de radiație are, într-adevăr, un nul în direcția $\theta = 60^\circ$ din care sosește semnalul de interferență și asigură eliminarea completă a acestuia din semnalul de ieșire. Șirul cu control adaptiv al lobului reușește să selecteze numai semnalul util pe baza diferenței dintre unghiurile de sosire dintre semnalul util și cel perturbator.

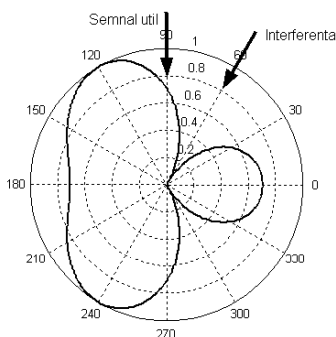


Figura nr. 1.7 – *Diagrama de radiație a șirului cu două elemente*

Exemplul de mai sus probează afirmația că valorile ponderilor pot fi optimizate pentru îndeplinirea funcției obiectiv prestabilite. Algoritmul de calcul s-a bazat pe cunoașterea direcțiilor din care sosesc semnalul util și cel de interferență. Se impune, însă, ca algoritmii de optimizare dezvoltati pentru situațiile reale să nu aibă nevoie de aceste informații deoarece ele nu sunt cunoscute. În plus, direcțiile semnalelor și numărul semnalelor de interferență sunt variabile în timp.

1.3 Criterii de optimizare

Fie un șir cu N antene echidistant așezate pe axa Oz a unui sistem cartezian de coordonate (figura nr. 1.8). Notăm cu Δ constanta de distanță a șirului, cu $x_m(t)$, $m = 1 \dots N$ – semnalul furnizat de antena m din șir, cu w_m – coeficientul complex (pondera) cu care este multiplicat semnalul x_m înainte de intrarea în sumator și cu $y(t)$ – semnalul de ieșire din sumator. Un procesor adaptiv compară semnalul $y(t)$ cu semnalul dorit $d^*(t)$ sau cu o replică a sa generată local și modifică ponderile individuale w_m , astfel ca semnalul $y(t)$ să se apropie cât mai mult de semnalul dorit. Prelucrarea erorii dintre cele două semnale se face într-o manieră specifică fiecărui algoritm în parte. (Asteriscul din notația $d^*(t)$ semnifică conjugare complexă și această reprezentare a fost aleasă pentru scrierea convenabilă a relațiilor ce vor interveni în continuare).

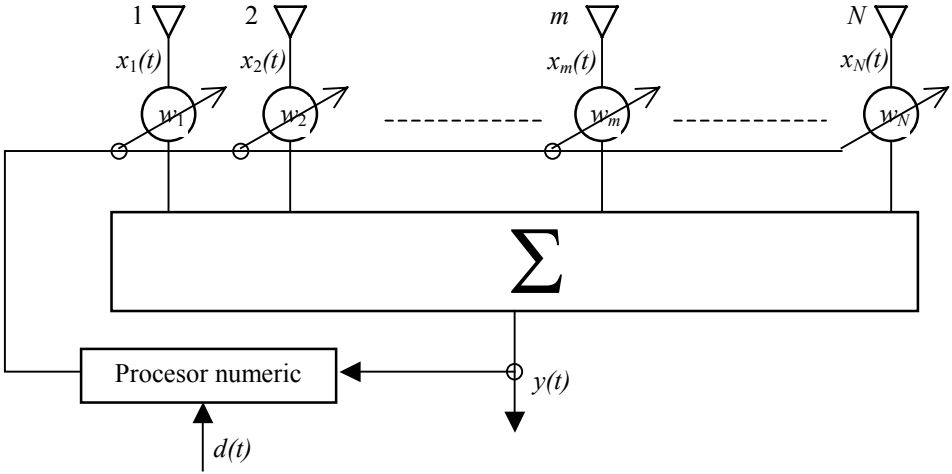


Figura nr. 1.8 – Arhitectura unui șir de antene cu formarea digitală adaptivă a lobului

Fiecare din semnalele $x_m(t)$ include o componentă de semnal util și o componentă de semnal perturbator (zgomot + interferență). Notăm cu N_p numărul total de semnale perturbatoare.

Contribuțiile semnalului util $s(t)$ la generarea fiecăruia dintre semnalele individuale $x_m(t)$ pot fi scrise compact sub forma:

$$\mathbf{s} = s(t)\mathbf{v}_u \quad (1.21)$$

iar contribuțiile semnalului perturbator $p_m(t)$ – sub forma:

$$\mathbf{p}_m = p_m(t)\mathbf{v}_{pm} \quad (1.22)$$

unde:

$$\mathbf{v}_u = \begin{bmatrix} 1 \\ e^{jk_0\Delta\cos\theta_u} \\ e^{j2k_0\Delta\cos\theta_u} \\ \vdots \\ e^{j(N-1)k_0\Delta\cos\theta_u} \end{bmatrix} \quad \mathbf{v}_{pm} = \begin{bmatrix} 1 \\ e^{jk_0\Delta\cos\theta_{pm}} \\ e^{j2k_0\Delta\cos\theta_{pm}} \\ \vdots \\ e^{j(N-1)k_0\Delta\cos\theta_{pm}} \end{bmatrix} \quad (1.23)$$

dacă θ_u și θ_{pm} sunt direcțiile din care provin semnalului util $s(t)$ și, respectiv, semnalul perturbator $p_m(t)$, iar $m = 1 \dots N_p$. Contribuția totală a semnalelor perturbatoare la generarea semnalelor antenelor individuale se scrie compact:

$$\mathbf{p} = \sum_{m=1}^{N_p} \mathbf{p}_m = \sum_{m=1}^{N_p} p_m(t) \mathbf{v}_{pm} \quad (1.24)$$

În concluzie:

$$\mathbf{x} = \mathbf{s} + \mathbf{p} = s(t) \mathbf{v}_u + \sum_{m=1}^{N_p} p_m(t) \mathbf{v}_{pm} \quad (1.25)$$

Pentru scrierea mai simplă a relațiilor ce vor urma se consideră că semnalele individuale ale antenelor se multiplică cu valorile conjugate ale ponderilor, astfel că:

$$y(t) = \mathbf{w}^H \cdot \mathbf{s} + \mathbf{w}^H \cdot \mathbf{p} \quad (1.26)$$

unde prin H s-a notat operația de transpunere hermitică (transpunere + conjugare complexă).

Minimizarea erorii pătratice medii (MMSE)

Definim eroarea dintre semnalul $y(t)$ furnizat de blocul de formare a lobului și cel dorit $d^*(t)$ ca fiind:

$$\varepsilon(t) = d^*(t) - y(t) = d^*(t) - \mathbf{w}^H \cdot \mathbf{x} \quad (1.27)$$

și impunem calculul ponderilor din condiția minimizării valorii medii a pătratului acestei erori, adică a expresiei:

$$\overline{\varepsilon^2(t)} = \overline{[d^*(t) - \mathbf{w}^H \cdot \mathbf{x}]^2} \quad (1.28)$$

prin barare superioară notându-se operatorul de mediere în timp aplicat expresiei respective.

Însă:

$$\begin{aligned} \varepsilon^2(t) &= d(t) \cdot d^*(t) - 2d^*(t) \mathbf{w}^H \cdot \mathbf{x} + (\mathbf{w}^H \cdot \mathbf{x}) \cdot (\mathbf{w}^H \cdot \mathbf{x})^H = \\ &= d^2(t) - 2d^*(t) \mathbf{w}^H \cdot \mathbf{x} + \mathbf{w}^H \cdot \mathbf{x} \cdot \mathbf{x}^H \cdot \mathbf{w} \end{aligned} \quad (1.29)$$

Ținând seama de faptul că ponderile nu sunt funcții de timp și, deci, vectorul ponderilor este invariant față de operatorul de mediere și notând:

$$\overline{d^*(t)\mathbf{x}} = \mathbf{r} \quad \text{și} \quad \overline{\mathbf{x} \cdot \mathbf{x}^H} = \mathbf{R} \quad (1.30)$$

unde \mathbf{R} este matricea de covarianță a semnalelor de la ieșirile antenelor individuale, obținem:

$$\overline{\varepsilon^2(t)} = \overline{d^2(t)} - 2\mathbf{w}^H \cdot \mathbf{r} + \mathbf{w}^H \cdot \mathbf{R} \cdot \mathbf{w} \quad (1.31)$$

Gradientul acestei expresii în raport cu vectorul \mathbf{w} al ponderilor este:

$$\nabla_{\mathbf{w}} \left(\overline{\varepsilon^2(t)} \right) = -2\mathbf{r} + 2\mathbf{R} \cdot \mathbf{w} \quad (1.32)$$

și din condiția ca el să fie nul (pentru ca eroarea pătratică medie să fie minimă) rezultă:

$$\mathbf{w}_{opt} = \mathbf{R}^{-1} \cdot \mathbf{r} \quad (1.33)$$

Recunoaștem în expresia de mai sus soluția Wiener optimă.

Dacă ținem seama de expresia (1.25) a vectorului \mathbf{x} și considerând că semnalele perturbatoare nu sunt corelate cu semnalul util obținem:

$$\mathbf{R} = \overline{(\mathbf{s} + \mathbf{p}) \cdot (\mathbf{s} + \mathbf{p})^H} = \overline{\mathbf{s} \cdot \mathbf{s}^H} + \overline{\mathbf{p} \cdot \mathbf{p}^H} = \mathbf{R}_u + \mathbf{R}_p = \overline{d^2(t)} \mathbf{v}_u \cdot \mathbf{v}_u^H + \mathbf{R}_p \quad (1.34)$$

unde $\mathbf{R}_u = \overline{\mathbf{s} \cdot \mathbf{s}^H}$ și $\mathbf{R}_p = \overline{\mathbf{p} \cdot \mathbf{p}^H}$ sunt matricile de covarianță ale componentelor de semnal util și, respectiv, de perturbație din semnalele de ieșire ale antenelor individuale.

Din expresia de mai sus a matricii de covarianță \mathbf{R} , potrivit identității lui Woodbury:

$$\mathbf{R}^{-1} = \frac{1}{1 + \overline{d^2(t)} \mathbf{v}_u^H \cdot \mathbf{R}_p^{-1} \cdot \mathbf{v}_u} \mathbf{R}_p^{-1} \quad (1.35)$$

astfel că vectorul ponderilor optime poate fi scris și ca:

$$\mathbf{w}_{opt} = \beta \mathbf{R}_p^{-1} \cdot \mathbf{v}_u \quad (1.36)$$

cu:

$$\beta = \frac{\overline{d^2(t)}}{1 + \overline{d^2(t)} \mathbf{v}_u^H \cdot \mathbf{R}_p^{-1} \cdot \mathbf{v}_u} \quad (1.37)$$

Maximizarea raportului semnal/zgomot (MSNR)

Presupunând cunoscute matricile \mathbf{R}_u și \mathbf{R}_p se pot scrie expresiile puterilor utilă și, respectiv, cea a perturbațiilor din semnalul de ieșire $y(t)$:

$$\sigma_s^2 = |\mathbf{w}^H \cdot \mathbf{s}|^2 = \mathbf{w}^H \cdot \mathbf{R}_u \cdot \mathbf{w} \quad \text{și} \quad \sigma_p^2 = |\mathbf{w}^H \cdot \mathbf{p}|^2 = \mathbf{w}^H \cdot \mathbf{R}_p \cdot \mathbf{w} \quad (1.38)$$

astfel că raportul semnal/(interferență + zgomot) la ieșirea din sumator este:

$$SINR = \frac{\sigma_s^2}{\sigma_p^2} = \frac{\mathbf{w}^H \cdot \mathbf{R}_u \cdot \mathbf{w}}{\mathbf{w}^H \cdot \mathbf{R}_p \cdot \mathbf{w}} \quad (1.39)$$

Anulând gradientul după vectorul \mathbf{w} al acestui raport se obține ecuația:

$$\mathbf{R}_u \cdot \mathbf{w} = \frac{\mathbf{w}^H \cdot \mathbf{R}_u \cdot \mathbf{w}}{\mathbf{w}^H \cdot \mathbf{R}_p \cdot \mathbf{w}} \mathbf{R}_p \cdot \mathbf{w} \quad \Leftrightarrow \quad \mathbf{R}_u \cdot \mathbf{w} = SINR \cdot \mathbf{R}_p \cdot \mathbf{w} \quad (1.40)$$

de unde:

$$\mathbf{R}_p^{-1} \cdot \mathbf{R}_u \cdot \mathbf{w} = SINR \cdot \mathbf{w} \quad (1.41)$$

adică valoarea maximă a $SINR$ este valoarea proprie maximă a matricii simetrice $\mathbf{R}_p^{-1} \cdot \mathbf{R}_u$.

Deci:

$$\begin{aligned} \mathbf{R}_u \cdot \mathbf{w}_{opt} &= SINR_{max} \cdot \mathbf{R}_p \cdot \mathbf{w}_{opt} \quad \Leftrightarrow \\ \Leftrightarrow \quad \overline{d^2(t)} \mathbf{v}_u \cdot \mathbf{v}_u^H \cdot \mathbf{w}_{opt} &= SINR_{max} \cdot \mathbf{R}_p \cdot \mathbf{w}_{opt} \end{aligned} \quad (1.42)$$

de unde:

$$\mathbf{w}_{opt} = \beta \mathbf{R}_p^{-1} \cdot \mathbf{v}_u \quad (1.43)$$

cu:

$$\beta = \frac{\overline{d^2(t)}}{SINR_{max}} \quad \mathbf{v}_u^H \cdot \mathbf{w}_{opt} \quad (1.44)$$

Se observă că soluția este, deasemenea, soluția Wiener optimă.

Minimizarea varianței semnalului de ieșire (MVDR)

Semnalul de ieșire din rețeaua de formare a lobului:

$$\mathbf{y} = \mathbf{w}^H \cdot \mathbf{x} = \mathbf{w}^H \cdot (\mathbf{s} + \mathbf{p}) = \mathbf{w}^H \cdot \mathbf{s} + \mathbf{w}^H \cdot \mathbf{p} \quad (1.45)$$

are o componentă de semnal util și o componentă dată de interferențe și zgomot (perturbații). Când nu se cunosc nici semnalul util și nici direcția din care el sosește o cale potrivită de a obține un optim la recepție este aceea de a minimiza componenta dată de perturbații. Pentru a ne asigura, totuși, că semnalul util are o valoare minimă acceptabilă putem impune, de exemplu, ca:

$$\mathbf{w}^H \cdot \mathbf{v}_u = g \quad (1.46)$$

unde g este o valoare ce ține seama de performanțele receptorului.

Dispersia semnalului de ieșire este:

$$\sigma_y^2 = \mathbf{w}^H \cdot \mathbf{R} \cdot \mathbf{w} = \mathbf{w}^H \cdot \mathbf{R}_u \cdot \mathbf{w} + \mathbf{w}^H \cdot \mathbf{R}_p \cdot \mathbf{w} \quad (1.47)$$

și problema de optimizare se reduce la găsirea vectorului \mathbf{w} care minimizează dispersia σ_y^2 cu constrângerea impusă de relația (1.46), ceea ce revine la minimizarea ultimului termen din membrul al doilea al relației (1.47).

Aplicăm metoda multiplicatorilor lui Lagrange și încercăm să aflăm minimul expresiei:

$$y'(t) = \mathbf{w}^H \cdot \mathbf{R}_p \cdot \mathbf{w} + 2\beta (1 - \mathbf{w}^H \cdot \mathbf{v}_u) \quad (1.48)$$

al cărui gradient în raport cu vectorul \mathbf{w} este:

$$\nabla_{\mathbf{w}}[y'(t)] = 2\mathbf{R}_p \cdot \mathbf{w} - 2\beta \mathbf{v}_u \quad (1.49)$$

Din condiția de anulare a acestei expresii rezultă:

$$\mathbf{w}_{opt} = \beta \mathbf{R}_p^{-1} \cdot \mathbf{v}_u \quad (1.50)$$

iar din relația de constrângere (1.46) rezultă:

$$\beta = \frac{g}{\mathbf{v}_u^H \mathbf{R}_p^{-1} \mathbf{v}_u} \quad (1.51)$$

Soluția (1.50) este, deasemenea, soluția Wiener optimă.

Dacă g este egal cu 1 se spune că răspunsul formatorului este *fără distorsiuni și cu dispersie minimă* (Minimum Variance Distortionless Response – de unde și acronimul MVDR asociat algoritmului de optimizare) sau, echivalent, cu asemănare maximă.

Valoarea maximă a SINR

Fiecare din criteriile de optimizare furnizează un vector al ponderilor \mathbf{w}_{opt} care satisface condiția de extrem specifică. Valoarea maximă corespunzătoare a SINR este:

$$\begin{aligned} SINR_{max} &= \frac{\mathbf{w}_{opt}^H \cdot \mathbf{R}_u \cdot \mathbf{w}_{opt}}{\mathbf{w}_{opt}^H \cdot \mathbf{R}_p \cdot \mathbf{w}_{opt}} = \frac{\left(\beta \mathbf{R}_p^{-1} \cdot \mathbf{v}_u \right)^H \cdot \left(\overline{d^2(t)} \mathbf{v}_u \cdot \mathbf{v}_u^H \right) \cdot \left(\beta \mathbf{R}_p^{-1} \cdot \mathbf{v}_u \right)}{\left(\beta \mathbf{R}_p^{-1} \cdot \mathbf{v}_u \right)^H \cdot \mathbf{R}_p \cdot \left(\beta \mathbf{R}_p^{-1} \cdot \mathbf{v}_u \right)} = \\ &= \overline{d^2(t)} \cdot \frac{\left(\mathbf{v}_u^H \cdot \mathbf{R}_p^{-1} \cdot \mathbf{v}_u \right) \left(\mathbf{v}_u^H \cdot \mathbf{R}_p^{-1} \cdot \mathbf{v}_u \right)}{\mathbf{v}_u^H \cdot \mathbf{R}_p^{-1} \cdot \mathbf{v}_u} = \overline{d^2(t)} \cdot \mathbf{v}_u^H \cdot \mathbf{R}_p^{-1} \cdot \mathbf{v}_u \end{aligned}$$

aceeași pentru toate criteriile !!!

1.4 Algoritmi de calcul adaptiv al ponderilor

Performanțele furnizate de cele trei criterii de optimizare sunt similare, astfel că pentru implementări practice se poate selecta oricare dintre ele. Însă tehnica concretă prin care se ajunge la valorile optime ale ponderilor trebuie selectată cu foarte mare atenție deoarece are implicații majore asupra performanțelor formatorului de lob. Toate tehnicile elaborate încearcă să realizeze un compromis acceptabil între viteza de convergență și complexitatea implementării hardware și software.

Algoritmul LMS (al celor mai mici pătrate)

Este unul dintre cei mai folosiți algoritmi și se bazează pe metoda gradientului (pantei maxime) în ipoteza rezonabilă că dacă se aplică vectorului ponderilor corecții succesive în vederea deplasării lui în sensul opus gradientului erorii se poate ajunge, în final, într-un punct de minim al acesteia. Concret, valoarea vectorului ponderilor la iterația curentă se calculează pe baza valorii lui la iterația precedentă și a gradientului erorii pătratice medii cu relația:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{1}{2} \mu \left[-\nabla_{\mathbf{w}} \overline{\varepsilon^2(t)} \right] \quad (1.53)$$

unde μ este un factor arbitrar ales prin care se poate ține sub control convergența algoritmului.

Ținând seama de expresia (1.32) a gradientului:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu [\mathbf{r} - \mathbf{R} \cdot \mathbf{w}(n)] \quad (1.54)$$

În această formă algoritmul de adaptare trebuie să dispună de informații privind covarianța semnalelor individuale furnizate de antene (\mathbf{R}) și despre parametrii semnalului util (\mathbf{r}), informații care nu sunt disponibile. Se utilizează, de aceea, valori aproximative ale acestora calculate pe baza celor mai recente estimări:

$$\hat{\mathbf{R}}(n) = \mathbf{x}(n) \cdot \mathbf{x}^H(n) \quad \text{și} \quad \hat{\mathbf{r}}(n) = \hat{d}^*(n) \mathbf{x}(n) \quad (1.55)$$

astfel că:

$$\begin{aligned} \hat{\mathbf{w}}(n+1) &= \\ &= \hat{\mathbf{w}}(n) + \mu \mathbf{x}(n) \left[\hat{d}^*(n) - \mathbf{x}^H(n) \cdot \hat{\mathbf{w}}(n) \right] = \hat{\mathbf{w}}(n) + \mu \mathbf{x}(n) \hat{\varepsilon}^*(n) \end{aligned} \quad (1.56)$$

Se observă că valorile ponderilor sunt modificate la fiecare nou eșantion al semnalelor de la antenele individuale ceea ce asigură algoritmului o foarte rapidă convergență dacă parametrii semnalelor nu se modifică sensibil în timp (quasi-staționaritate).

Algoritmul LMS este extrem de simplu de implementat, însă are o slabă viteză de convergență atunci când matricea de covarianță \mathbf{R} are valori proprii mult diferite între ele.

Algoritmul SMI (inversarea matricii eșantioanelor)

Valorile optime ale ponderilor s-ar putea determina direct din relația (1.33) dacă s-ar cunoaște matricea de covarianță \mathbf{R} și vectorul \mathbf{r} . Din păcate acești parametri nu se cunosc și, în plus, ei se modifică în permanență. Relația (1.33) încă se mai poate folosi dacă se utilizează estimări ale matricii \mathbf{R} și vectorului \mathbf{r} , estimări bazate pe valorile eșantioanelor lor pe un interval de timp convenabil ales. Dacă, de exemplu, N_{max} și N_{min} sunt limitele intervalului de observație (fereastra), atunci valorile estimate căutate sunt:

$$\hat{\mathbf{R}} = \sum_{N_{min}}^{N_{max}} \mathbf{x}(i) \cdot \mathbf{x}^H(i) \quad \text{și} \quad \hat{\mathbf{r}} = \sum_{N_{min}}^{N_{max}} d^*(i) \mathbf{x}(i) \quad (1.57)$$

În consecință:

$$\hat{\mathbf{w}} = \hat{\mathbf{R}}^{-1} \cdot \hat{\mathbf{r}} \quad (1.58)$$

ceea ce conduce la o eroare de estimare:

$$e = \hat{\mathbf{R}} \cdot \mathbf{w}_{opt} - \hat{\mathbf{r}} \quad (1.59)$$

Algoritmul SMI are, teoretic, o viteză mai mare de convergență decât algoritmul LMS, însă prezintă două mari dezavantaje: (a) *complexitate mare* de implementare chiar și pentru circuitele VLSI moderne și (b) *instabilitate numerică* datorită preciziei finite de reprezentare a numerelor în procesoarele numerice și a volumului uriaș de calcule necesar inversării unei matrici de mari dimensiuni.

Algoritmul RLS (LMS recursiv)

În locul estimării matricii \mathbf{R} și a vectorului \mathbf{r} pe baza celor mai recente eșantioane ale lor se pot folosi relațiile:

$$\begin{aligned} \hat{\mathbf{R}}(n) &= \gamma \hat{\mathbf{R}}(n-1) + \mathbf{x}(n) \cdot \mathbf{x}^H(n) \\ \hat{\mathbf{r}}(n) &= \gamma \hat{\mathbf{r}}(n-1) + d^*(n) \mathbf{x}^H(n) \end{aligned} \quad (1.60)$$

unde γ este un coeficient strict pozitiv și subunitar denumit „factor de uitare”. Prin alegerea valorii lui se selectează ponderea pe care o au eșantioanele anterioare în estimarea valorilor curente ale parametrilor \mathbf{R} și \mathbf{r} .

Relația (1.60) permite obținerea unei expresii a matricii inverse \mathbf{R}^{-1} sub forma (identitatea lui Woodbury):

$$\mathbf{R}^{-1}(n) = \gamma^{-1} \left[\mathbf{R}^{-1}(n-1) - \mathbf{q}(n) \cdot \mathbf{x}(n) \cdot \mathbf{R}^{-1}(n-1) \right] \quad (1.61)$$

unde $\mathbf{q}(n)$ este un vector de câștig ce are expresia:

$$\mathbf{q}(n) = \frac{\gamma^{-1} \mathbf{R}^{-1}(n-1) \cdot \mathbf{x}(n)}{1 + \gamma^{-1} \mathbf{x}^H(n) \cdot \mathbf{R}^{-1}(n-1) \cdot \mathbf{x}(n)} \quad (1.62)$$

astfel că:

$$\begin{aligned} \hat{\mathbf{w}}(n) &= \mathbf{R}^{-1}(n) \cdot \mathbf{r}(n) = \\ &= \gamma^{-1} \left[\mathbf{R}^{-1}(n-1) - \mathbf{q}(n) \cdot \mathbf{x}(n) \cdot \mathbf{R}^{-1}(n-1) \right] \left[\gamma \hat{\mathbf{r}}(n-1) + d^*(n) \mathbf{x}^H(n) \right] \end{aligned} \quad (1.63)$$

sau:

$$\hat{\mathbf{w}}(n) = \hat{\mathbf{w}}(n-1) + \left[d^*(n) - \mathbf{w}^H(n-1) \cdot \mathbf{x}(n) \right] \mathbf{q}(n) \quad (1.64)$$

Se observă că operația de inversare a matricii de covarianță este înlocuită la fiecare iterație cu o operație de divizare ceea ce conduce la creșterea cu un ordin de mărime a vitezei de convergență.

Algoritm bazat pe rețele neurale artificiale

Interesul manifestat pentru utilizarea rețelelor neurale artificiale (RNA) în soluționarea unor probleme de optimizare se datorează faptului că acestea sunt alcătuite dintr-un număr foarte mare de elemente simple (neuroni) interconectate între ele și capabile să proceseze în paralel informația de intrare. Ca urmare, RNA se pot adapta unor medii complexe și pot răspunde rapid la schimbările petrecute în acestea.

Într-o RNA (figura 1.9a) cu N neuroni fiecare neuron are $N+1$ intrări și o ieșire. Pe intrările unui neuron se aplică semnalele de ieșire de la ceilalți neuroni din rețea (interconectarea), un semnal de la propria ieșire (reacția) și un semnal extern. Suma ponderată a celor $N+1$ semnale de intrare reprezintă *potențialul intern* al neuronului. Semnalul de ieșire se calculează pe baza potențialului intern prin intermediul unei funcții neliniare denumită *funcție sigmoidală* (figura 1.9b). Principalele caracteristici ale unei funcții neliniare utilizabilă ca funcție sigmoidală sunt (a) panta mare de

variație în jurul originii și (b) limitarea ieșirii pentru valori mult diferite de 0 ale intrării. Ieșirea unei RNA este dată de vectorul semnalelor de ieșire ale neuronilor componenți (sau numai ieșirile neuronilor din stratul de ieșire dacă aceștia sunt organizați pe straturi).

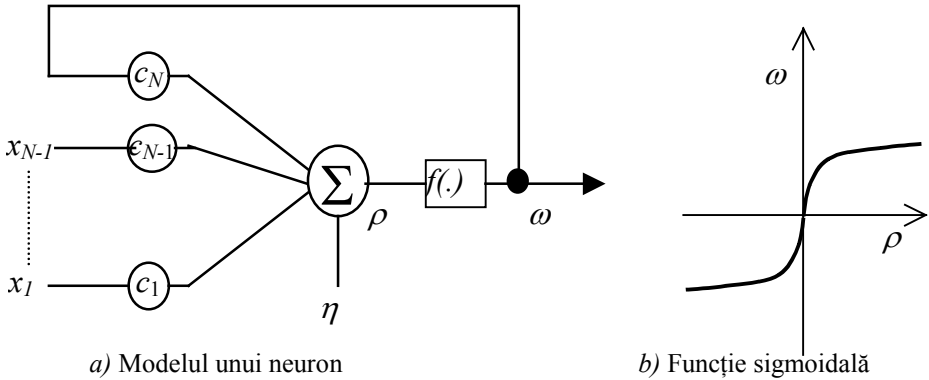


Figura nr. 1.9 – Modelul unui neuron și exemplu de funcție sigmoidală

Potrivit notațiilor din figura 1.9:

$$\rho = \sum_{i=1}^{N-1} c_i x_i + c_N \omega + \eta \quad \omega = f(\rho) \quad (1.65)$$

Într-o RNA de tip Hopfield ponderile interconexiunilor dintre neuroni sunt simetrice (adică ponderea cu care este sumată ieșirea neuronului p la intrarea neuronului q este aceeași cu ponderea cu care este sumată ieșirea neuronului q la intrarea neuronului p) și nu există reacție de la ieșirea unui neuron către propria intrare.

Pentru o RNA Hopfield cu funcționare discretă putem scrie că:

$$\rho_i(n+1) = \sum_{\substack{j=1 \\ j \neq i}}^N c_{ij}(n) \omega_j(n) + \eta_i(n) \quad (1.66)$$

$$\omega_i(n+1) = f[\rho_i(n+1)], \quad i = 1 \dots N$$

iar dacă introducem următoarele notații:

$$\mathbf{p} = \begin{bmatrix} \rho_1 \\ \vdots \\ \rho_N \end{bmatrix}, \quad \mathbf{\eta} = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_N \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} c_{11} & \dots & c_{1N} \\ \vdots & \ddots & \vdots \\ c_{N1} & \dots & c_{NN} \end{bmatrix}, \quad \mathbf{\omega} = \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_N \end{bmatrix} \quad (1.67)$$

atunci ieșirea RNA se poate scrie compact sub forma:

$$\mathbf{\omega}(n+1) = f[\mathbf{p}(n+1)] \quad \text{cu} \quad \mathbf{p}(n+1) = \mathbf{C}(n) \cdot \mathbf{\omega}(n) + \mathbf{\eta}(n) \quad (1.68)$$

Dacă definim energia unei RNA ca fiind:

$$\mathbf{E} = -\frac{1}{2} \sum_i \sum_j c_{ij} \omega_i \omega_j - \sum_i \eta_i \omega_i \quad (1.69)$$

sau, compact:

$$\mathbf{E} = -\frac{1}{2} \mathbf{\omega}^T \cdot \mathbf{C} \cdot \mathbf{\omega} - \mathbf{\omega}^T \cdot \mathbf{\eta} \quad (1.70)$$

se poate arăta că ea este o funcție de tip Liapunov, adică o funcție pozitiv definită și a cărei valoare nu crește la modificarea argumentului (vectorul $\mathbf{\omega}$, în acest caz). Aceasta înseamnă că la modificarea ponderilor interconexiunilor dintre neuroni energia RNA scade permanent până când atinge minimul, după care rămâne nemodificată. Pe această proprietate se bazează utilizarea RNA în soluționarea unor probleme de optimizare, probleme care constau în cele mai multe cazuri în găsirea minimului unei funcții de cost.

În cazul concret al formării lobului unui șir de antene, dacă punem în evidență părțile reală și imaginară ale semnalelor $d(t) = d_R(t) + j d_I(t)$ și $y(t) = y_R(t) + j y_I(t)$, precum și ale vectorilor \mathbf{x} , \mathbf{w} și definim:

$$\mathbf{W} = \begin{bmatrix} \mathbf{w}_R \\ \mathbf{w}_I \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} \mathbf{x}_R & \mathbf{x}_I \\ \mathbf{x}_R & -\mathbf{x}_I \end{bmatrix} \quad (1.71)$$

putem scrie că:

$$\begin{bmatrix} y_R & y_I \end{bmatrix} = \mathbf{W}^T \cdot \mathbf{X} \quad (1.72)$$

iar eroarea pătratică medie devine:

$$\overline{\varepsilon^2} = \overline{(d_R - y_R)^2 + (d_I - y_I)^2} = \overline{d_R^2 + d_I^2 + y_R^2 + y_I^2 - 2(d_R y_R + d_I y_I)} \quad (1.73)$$

Deoarece componentele semnalului dorit $d(t)$ sunt constante, minimizarea erorii pătratice medii revine la a minimiza funcția:

$$\begin{aligned} \psi &= \overline{y_R^2 + y_I^2 - 2(d_R y_R + d_I y_I)} = -2[y_R \ y_I] \cdot \begin{bmatrix} d_R \\ d_I \end{bmatrix} + [y_R \ y_I] \cdot \begin{bmatrix} y_R \\ y_I \end{bmatrix} = \\ &= \overline{-2 \mathbf{W}^T \cdot \mathbf{X} \cdot \begin{bmatrix} d_R \\ d_I \end{bmatrix} + \mathbf{W}^T \cdot \mathbf{X} \cdot \mathbf{X}^T \cdot \mathbf{W}} \end{aligned} \quad (1.74)$$

Expresia ψ devine identică cu energia unei RNA de tip Hopfield dacă facem echivalențele:

$$\mathbf{C} = -2 \overline{\mathbf{X} \cdot \mathbf{X}^T} \quad \boldsymbol{\eta} = 2 \overline{\mathbf{X} \cdot \begin{bmatrix} d_R \\ d_I \end{bmatrix}} \quad \boldsymbol{\omega} = \mathbf{W} \quad (1.75)$$

și, deci, i se poate determina minimul cu ajutorul rețelei Hopfield echivalente.