

Hírközléelmélet I.

Dr. Pap László

2009. április 10.

Tartalomjegyzék

1. A modulált jelek általános leírása	5
1.1. A modulált jelek vektortérbeli leírása	5
1.2. A fehér Gauss-zaj leírása a vektortérben	7
1.3. Példák a modulált jelek vektortérbeli leírására	9
2. Az optimális koherens vevő	19
2.1. Az optimális demodulálási szabály	19
2.2. Példák a döntési tartományok kiszámítására	20
2.3. Az optimális koherens vevő felépítése	22
2.4. Az optimális koherens vevő hibavalószínűségeinek meghatározása	23
2.5. Példák a hibavalószínűség számítására	26
2.6. A hibaarány közelítő számítása	28
2.7. A koherens modulációs rendszerek általános jellemzése	31
2.8. Példák a koherens modulációs rendszerek általános jellemzésére	32
3. Az optimális nem koherens vevő struktúrája és hibaaránya	39
3.1. A nem koherens jelek jeltérbeli ábrázolása	39
3.2. Optimális nem koherens vevőstruktúrák	43
3.3. A nem koherens optimális vevők működésének illusztratív összehasonlítása	46
3.4. A nem koherens rendszerek hibavalószínűsége	48
3.5. Példa a nem koherens rendszerek hibaarányának számítására	51
4. A koherens és nem koherens átviteli rendszerek összehasonlítása	55
4.1. Koherens csatorna	55
4.2. Nem koherens csatorna	56
5. A vektortér kiterjesztése a sávhatárolt jelekre	59
5.1. A négyzetesen integrálható függvények L_2 tere	60
5.2. A jeltér általános definíciója	61
5.3. Az alapsávi sávkorlátozott jelek leírása a jeltérben	61
5.4. Az áteresztő sávi sávkorlátozott jelek leírása a jeltérben	63
5.5. A fehér Gauss-zaj leírása az általános jeltérben	65
5.6. Az áteresztő sávi sávhatárolt jelek leírása a komplex számok terében	67
5.7. Az ortogonális PAM és QAM modulációk	68
5.8. A kódolatlan rendszerek teljesítőképessége a Shannon-kapacitáshoz viszonyítva	73
5.9. Az M -PAM és az $(M \times M)$ -QAM teljesítőképessége	75

6. Kis jelterek teljesítőképessége	81
6.1. Jelkonstellációk fehér Gauss-zajos csatorna esetén	81
6.2. A teljesítőképesség vizsgálata a teljesítménykorlátozott tartományban	85
6.3. A teljesítőképesség vizsgálata a sávkorlátozott tartományban	90
6.4. Összefoglalás	91
7. Bevezetés a bináris kódok világába	93
7.1. Bináris jelkonstellációk	93
7.2. A bináris lineáris blokk kódok, mint bináris vektorterek	94
7.3. A lineáris blokk kódok az euklideszi térben	98
7.4. Reed-Muller kódok	100
7.5. A bináris blokk kódok dekódolása	106
8. A modulált jelek spektrális vizsgálata	113
8.1. A ciklostacionárius jelek tulajdonságai	113
8.2. A véletlen fázisú szinuszos jel teljesítménysűrűség-függvénye	115
8.3. Az alapsávi PAM jelek teljesítménysűrűség-függvénye	116
8.4. Illusztratív példák a PAM jelek spektrális analízisére	117
8.5. Az általános optimális PAM rendszer vizsgálata	118
8.6. Részleges válaszfüggvényű PAM típusú rendszerek	120
8.7. Példák a részleges válaszfüggvényű rendszerek spektrális vizsgálatára	121
8.8. Általános modulációs rendszer vizsgálata	123
8.9. A folytonos fázisú FM modulált jelek spektruma	130

1. fejezet

A modulált jelek általános leírása

1.1. A modulált jelek vektortérbeli leírása

Legyen a forrás által előállított szimbólumok sorozata $\{\xi_{k_i}\}$, amelyre igaz, hogy minden eleme egy M értékészletű ábécéből veszi fel az értékeit, azaz $k_i \in \{1, \dots, k, \dots, M\}$, és az i index az időbeli sorrendet jelöli. (Ha az eredeti forrás L értékészletű szimbólumsorozatot állít elő, és azokból K darabot összefogva állítjuk elő a ξ szimbólumokat, akkor $M = L^K$.) Jelöljük a szimbólumidőt T -vel, így módon az i -edik időrészben a ξ_{k_i} szimbólumot vesszük át, azaz, ha $k_i = h$, akkor a h -adik szimbólumot.

A $\{\xi_{k_i}\}$ szimbólumsorozat átviteléhez időben folytonos függvényeket használunk, ezeket $\{x_{k_i}(t)\}$ -vel jelöljük. A moduláció tehát egyszerű leképezés, miszerint a ξ_{k_i} szimbólumhoz az $x_{k_i}(t - iT)$ időben eltolts elemi jelet rendeljük hozzá.

Tételezzük fel, hogy az elemi jel csak a $[0, T)$ tartományban különbözik nullától, azaz

$$x_k(t) = \begin{cases} \text{tetszőleges,} & t \in [0, T) \\ 0, & t \notin [0, T) \end{cases},$$

vagyis az elemi jelek tartója a $[0, T)$ tartomány. (Természetesen ez a megkötés nem teljesül általában, de a tárgyalás egyszerűsítése érdekében célszerű bevezetni, így módon ugyanis az egyes időrészekben a jelek függetlenül kezelhetők, épp ezért a továbbiakban az i időindexet elhagyva csak ξ_k és $x_k(t)$ jelöléseket alkalmazunk.)

Vezessük be a $\{\varphi_j(t)\}$ ortonormált bázist az alábbi tulajdonságokkal:

- $\varphi_j(t)$ tartója a $[0, T)$ tartomány, $j = 1, \dots, N$;
- a skaláris szorzat

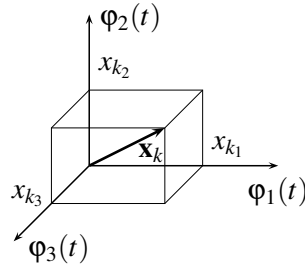
$$\langle \varphi_j(t), \varphi_k(t) \rangle = \int_0^T \varphi_j(t) \varphi_k(t) dt = \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases}$$

- a $\{\varphi_j(t)\}$ bázis teljes az $\{x_k(t)\}$ elemi jelekre, ami annyit jelent, hogy érvényesek az egyenletek

$$\left. \begin{aligned} x_k(t) &= \sum_{j=1}^N x_{k_j} \varphi_j(t) \\ x_{k_j} &= \int_0^T x_k(t) \varphi_j(t) dt \end{aligned} \right\} \begin{array}{l} \text{általánosított} \\ \text{Fourier-sorfejtés} \end{array}$$

Megjegyzendő, hogy az ortonormáltság az ún. Gram–Schmidt-ortogonalizálással mindig biztosítható, azaz:

1. $x_1(t) \stackrel{!}{=} a \varphi_1(t)$



1.1. ábra. Jelek vektoros ábrázolása

$$\varphi_1(t) = \frac{x_1(t)}{a}, \int_0^T \varphi_1^2(t) dt = 1 = \frac{1}{a^2} \int_0^T x_1^2(t) dt$$

$$a^2 = \int_0^T x_1^2(t) dt$$

$$2. x_2(t) \stackrel{!}{=} b\varphi_1(t) + c\varphi_2(t)$$

$\int_0^T x_2(t)\varphi_1(t) dt = b$, így $x_2(t) - b\varphi_1(t) = c\varphi_2(t)$, ahol a bal oldal ismert, tehát $\varphi_2(t)$ és c az 1. pontbeli eljárással meghatározható.

3. Ezután az eljárás folytatható.

Mindebből nyilvánvaló, hogy a teljes tér dimenziója $N \leq M$, ahol N az ortonormált bázis elemeinek a száma, M az elemi jelek (üzenetek) száma.

A jelek tehát kétféle módon ábrázolhatók:

- $x_k(t)$ időfüggvényekkel;
- $\mathbf{x}_k = \{x_{k_1}, \dots, x_{k_j}, \dots, x_{k_N}\}$ vektorokkal;
- Illusztráció: 1.1. ábra ($N = 3$).

A jel energiájának kiszámítása:

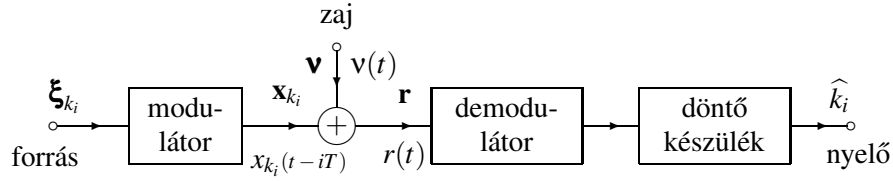
$$\begin{aligned} E_k &= \int_0^T x_k^2(t) dt = \\ &= \int_0^T \left(\sum_{j=1}^N x_{k_j} \varphi_j(t) \right) \left(\sum_{l=1}^N x_{k_l} \varphi_l(t) \right) dt = \\ &= \sum_{j=1}^N \sum_{l=1}^N x_{k_j} x_{k_l} \int_0^T \varphi_j(t) \varphi_l(t) dt = \\ &= \sum_{j=1}^N x_{k_j}^2 \end{aligned}$$

A jel energiája az \mathbf{x}_k vektor hosszának a négyzete.

A moduláció tehát nem jelent mást, mint a szimbólumok hozzárendelését időben folytonos jelekhez, vagy N dimenziós térbeli vektorokhoz, tehát:

$$\xi_{k_i} \implies x_{k_i}(t - iT) \quad \text{vagy} \quad \xi_{k_i} \implies \mathbf{x}_{k_i}.$$

A vizsgált csatorna az 1.2. ábrán látható. A csatorna ún. additív zajos csatorna. Az üzenetek visszaállítása nem jelent mást, mint minden időrészben a k_i becslését, azaz \hat{k}_i meghatározását. A továbbiakban a $v(t)$ zajról azt feltételezzük, hogy az fehér Gauss-zaj.



1.2. ábra. A csatorna modellje

1.2. A fehér Gauss-zaj leírása a vektortérben

A rendszer egységes kezelése érdekében írjuk le a fehér Gauss-zajt is az $\{x_k(t)\}$ által generált vektortérben. A fehér Gauss-zaj jellemző paraméterei:

- N_0 , egyoldalas teljesítménysűrűség $\left[\frac{W}{Hz}\right]$;
- $\frac{N_0}{2}$, kétoldalas teljesítménysűrűség $\left[\frac{W}{Hz}\right]$;
- $s_0 = \frac{1}{2\pi} \frac{N_0}{2}$, kétoldalas teljesítménysűrűség $\left[\frac{W}{\text{rad/sec}}\right]$.

A fehér Gauss-zaj autokorrelációs függvénye

$$R(\tau) = \int_{-\infty}^{+\infty} e^{j\omega\tau} s(\omega) d\omega = \frac{N_0}{2} \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{j\omega\tau} d\omega = \frac{N_0}{2} \delta(\tau)$$

Definiáljuk a \mathbf{v} zajvektort az N dimenziós térben

$$\mathbf{v}_j = \int_0^T \mathbf{v}(t) \varphi_j(t) dt; \quad \mathbf{v} = \{v_1, \dots, v_j, \dots, v_N\}$$

Mivel $\mathbf{v}(t)$ sztochasztikus folyamat, \mathbf{v} valószínűségi vektorváltozó, melynek elemei az alábbi tulajdonságokkal jellemezhetők:

- v_j , Gauss-eloszlású valószínűségi változó, mivel $\mathbf{v}(t)$ -ből lineáris transzformációval állítottuk elő.
- $\mathbf{E}\{v_j\} = 0$, ugyanis

$$\mathbf{E}\left\{\int_0^T \mathbf{v}(t) \varphi_j(t) dt\right\} = \int_0^T \mathbf{E}\{\mathbf{v}(t)\} \varphi_j(t) dt = 0$$

- korreláció

$$\begin{aligned} \mathbf{E}\{v_j v_k\} &= \mathbf{E}\left\{\left(\int_0^T \mathbf{v}(t) \varphi_j(t) dt\right) \left(\int_0^T \mathbf{v}(\rho) \varphi_k(\rho) d\rho\right)\right\} = \\ &= \mathbf{E}\left\{\int_0^T \int_0^T \mathbf{v}(t) \mathbf{v}(\rho) \varphi_j(t) \varphi_k(\rho) dt d\rho\right\} = \\ &= \int_0^T \int_0^T \mathbf{E}\{\mathbf{v}(t) \mathbf{v}(\rho)\} \varphi_j(t) \varphi_k(\rho) dt d\rho = \\ &= \frac{N_0}{2} \int_0^T \int_0^T \delta(t - \rho) \varphi_j(t) \varphi_k(\rho) dt d\rho = \end{aligned}$$

$$\begin{aligned}
&= \frac{N_0}{2} \int_0^T \varphi_j(t) \varphi_k(t) dt = \\
&= \begin{cases} \frac{N_0}{2}, & j = k \\ 0, & j \neq k \end{cases}
\end{aligned}$$

- Az egyes időrésekben a \mathbf{v} vektorok függetlenek egymástól, így a \mathbf{v} vektor független Gauss-eloszlású valószínűségi vektorváltozó az alábbi sűrűségfüggvénnyel:

$$p_{\mathbf{v}}(\mathbf{x}) = \left(\frac{1}{\sqrt{2\pi\sigma}} \right)^N \prod_{j=1}^N \exp\left(-\frac{x_j^2}{2\sigma^2}\right) = \left(\frac{1}{\sqrt{\pi N_0}} \right)^N \prod_{j=1}^N \exp\left(-\frac{x_j^2}{N_0}\right)$$

Fontos megjegyezni, hogy

$$\mathbf{v}(t) \neq \sum_{j=1}^N v_j \varphi_j(t),$$

tehát a $\{\varphi_j(t)\}$ bázis a fehér Gauss-zajra nem teljes.

$$\mathbf{v}(t) = \sum_{j=1}^N v_j \varphi_j(t) + \tilde{\mathbf{v}}(t),$$

ahol $\tilde{\mathbf{v}}(t)$ a fehér Gauss-zaj azon összetevőinek az összessége, amelyek a hasznos jel által generált (kifeszített) téren kívül esnek.

A vevő bemenetére jutó jel az

$$r(t) = x_k(t) + \mathbf{v}(t),$$

feltéve, ha tudjuk, hogy a vizsgált időrészben a k -adik szimbólumot állította elő a forrás. A vevő bemenetére jutó jel leírható az általunk definiált vektortérben is, így létrejön az $\mathbf{r} = \{r_1, \dots, r_j, \dots, r_N\}$ vektor, ahol

$$r_j = \int_0^T r(t) \varphi_j(t) dt = x_{k_j} + v_j.$$

A korábbi vizsgálatból azonban nyilvánvaló, hogy

$$r(t) \neq \sum_{j=1}^N r_j \varphi_j(t),$$

hiszen

$$r(t) = \sum_{j=1}^N x_{k_j} \varphi_j(t) + \sum_{j=1}^N v_j \varphi_j(t) + \tilde{\mathbf{v}}(t).$$

Felvetődik tehát a kérdés, hogy az $r(t)$ vételéhez, pontosabban az $x_k(t)$ megfigyeléséhez elegendő-e csak az \mathbf{r} vektor elemeit meghatározni? A válasz erre a kérdésre igen, vagyis \mathbf{r} megfigyelése *elégseges statisztika* az üzenet vételéhez.

Érdemes megjegyezni, hogy

$$\begin{aligned}
\mathbf{E} \{r_j \tilde{\mathbf{v}}(t)\} &= \mathbf{E} \left\{ \tilde{\mathbf{v}}(t) \int_0^T r(\rho) \varphi_j(\rho) d\rho \right\} = \\
&= \mathbf{E} \left\{ \tilde{\mathbf{v}}(t) \int_0^T (x_k(\rho) + \mathbf{v}(\rho)) \varphi_j(\rho) d\rho \right\} =
\end{aligned}$$

$$\begin{aligned}
&= \mathbf{E} \left\{ \left(\mathbf{v}(t) - \sum_{l=1}^N \mathbf{v}_l \varphi_l(t) \right) \int_0^T (x_k(\rho) + \mathbf{v}(\rho)) \varphi_j(\rho) d\rho \right\} = \\
&= \mathbf{E} \left\{ \left(\mathbf{v}(t) - \sum_{l=1}^N \mathbf{v}_l \varphi_l(t) \right) (x_{k_j} + \mathbf{v}_j) \right\} = \\
&= \mathbf{E} \left\{ \left(\mathbf{v}(t) - \sum_{l=1}^N \mathbf{v}_l \varphi_l(t) \right) \mathbf{v}_j \right\} = \\
&= \mathbf{E} \{ \mathbf{v}(t) \mathbf{v}_j \} - \sum_{l=1}^N \mathbf{E} \{ \mathbf{v}_l \mathbf{v}_j \} \varphi_l(t) = \\
&= \mathbf{E} \left\{ \mathbf{v}(t) \int_0^T \mathbf{v}(\rho) \varphi_j(\rho) d\rho \right\} - \frac{N_0}{2} \varphi_j(t) = \\
&= \int_0^T \mathbf{E} \{ \mathbf{v}(\rho) \mathbf{v}(t) \} \varphi_j(\rho) d\rho - \frac{N_0}{2} \varphi_j(t) = \\
&= \frac{N_0}{2} \int_0^T \delta(t - \rho) \varphi_j(\rho) d\rho - \frac{N_0}{2} \varphi_j(t) = 0,
\end{aligned}$$

azaz a $\tilde{\mathbf{v}}(t)$ független az \mathbf{r} elemeitől, így nem hordoz semmiféle információt az \mathbf{x}_k vektorról.

A vevőbe jutó jelből származó \mathbf{r} vektor statisztikája a $p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k)$ feltételes valószínűségi sűrűségfüggvénnyel adható meg. Ennek előállításához tudjuk:

- Az \mathbf{r} vektor feltételesen Gauss-eloszlású, feltéve azt, hogy a k -adik üzenetet küldték.
- Az \mathbf{r} vektor elemei feltételesen függetlenek egymástól.
- $\mathbf{E} \{ r_j | \mathbf{x}_k \} = \mathbf{E} \left\{ \int_0^T r(t) \varphi_j(t) dt \right\} = \mathbf{E} \left\{ \int_0^T (x_k(t) + \mathbf{v}(t)) \varphi_j(t) dt \right\} = x_{k_j}$
- Az r_j szórásnégyzete az alábbi kifejezéssel határozható meg:

$$\mathbf{E} \{ (r_j - x_{k_j})(r_l - x_{k_l}) | \mathbf{x}_k \} = \mathbf{E} \{ \mathbf{v}_j \mathbf{v}_l \} = \begin{cases} \frac{N_0}{2}, & j = l \\ 0, & j \neq l \end{cases}$$

Ezért az \mathbf{r} vektor elemei feltételesen független Gauss-eloszlású valószínűségi változók. Az \mathbf{r} elemeinek feltételes együttes eloszlása

$$p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k) = \prod_{j=1}^N \frac{1}{\sqrt{\pi N_0}} \exp \left(-\frac{(y_j - x_{k_j})^2}{N_0} \right) = \prod_{j=1}^N p_{r_j}(y_j | x_{k_j}),$$

ahol

$$p_{r_j}(y_j | x_{k_j}) = \frac{1}{\sqrt{\pi N_0}} \exp \left(-\frac{(y_j - x_{k_j})^2}{N_0} \right).$$

1.3. Példák a modulált jelek vektortérbeli leírására

QPSK jel

$M = 4$, elemi jelek:

$$\begin{aligned} x_1(t) &= \cos(\omega_0 t) = -x_2(t) \\ x_3(t) &= \sin(\omega_0 t) = -x_4(t) \end{aligned} \quad t \in [0, T)$$

Az ortonormált bázisfüggvények előállítására:

$$x_1(t) = a\varphi_1(t),$$

$$\begin{aligned} a^2 &= \int_0^T x_1^2(t) dt = \\ &= \int_0^T \cos^2(\omega_0 t) dt = \\ &= \frac{1}{2} \int_0^T (1 + \cos(2\omega_0 t)) dt = \\ &= \frac{T}{2} + \frac{1}{2} \left[\frac{\sin(2\omega_0 t)}{2\omega_0} \right]_0^T = \frac{T}{2}, \end{aligned}$$

ha $2\omega_0 T = \pi + n\pi$. Ebből

$$\varphi_1(t) = \sqrt{\frac{2}{T}} \cos(\omega_0 t), \quad x_{11} = \sqrt{\frac{T}{2}}, \quad x_{12} = 0,$$

és értelemszerűen

$$x_{21} = -\sqrt{\frac{T}{2}}, \quad x_{22} = 0.$$

A Gram–Schmidt-ortogonalizálás lépéseit követve:

$$x_3(t) = b\varphi_1(t) + c\varphi_2(t),$$

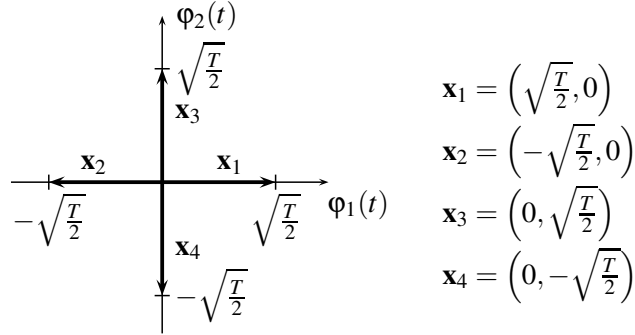
$$\begin{aligned} b &= \int_0^T x_3(t)\varphi_1(t) dt = \\ &= \sqrt{\frac{2}{T}} \int_0^T \cos(\omega_0 t) \sin(\omega_0 t) dt = \\ &= \sqrt{\frac{2}{T}} \frac{1}{2} \int_0^T \sin(2\omega_0 t) dt = \\ &= \sqrt{\frac{2}{T}} \frac{1}{2} \left[-\frac{\cos(2\omega_0 t)}{2\omega_0} \right]_0^T = 0, \end{aligned}$$

ha $2\omega_0 T = 2\pi + n2\pi$. Ezért $x_3(t) = c\varphi_2(t)$, amiből az előző lépésekhez hasonlóan

$$\varphi_2(t) = \sqrt{\frac{2}{T}} \sin(\omega_0 t), \quad x_{31} = 0, \quad x_{32} = \sqrt{\frac{T}{2}},$$

és értelemszerűen

$$x_{41} = 0, \quad x_{42} = -\sqrt{\frac{T}{2}}.$$



1.3. ábra. A QPSK jel vektoros ábrázolása

A vektorokat egy kétdimenziós térben lehet ábrázolni (1.3. ábra), tehát $N = 2 < M$. Az ábrázolás a szokásos, ún. konstellációs diagramra emlékeztet, tartalmilag azonban teljesen más értelmet hordoz. Érdeemes azt is megemlíteni, hogy az ábráról a vektorok közötti távolság is leolvasható:

$$\begin{aligned}
 \|\mathbf{x}_1 - \mathbf{x}_3\|^2 &= \|x_1(t) - x_3(t)\|^2 = \\
 &= \int_0^{T_s} (\cos(\omega_0 t) - \sin(\omega_0 t))^2 dt = \\
 &= (x_{11} - x_{31})^2 + (x_{12} - x_{32})^2 = \\
 &= \left(\sqrt{\frac{T}{2}}\right)^2 + \left(\sqrt{\frac{T}{2}}\right)^2 = T
 \end{aligned}$$

$$\begin{aligned}
 \|\mathbf{x}_1 - \mathbf{x}_2\|^2 &= \|x_1(t) - x_2(t)\|^2 = \\
 &= \int_0^{T_s} (2\cos(\omega_0 t))^2 dt = \\
 &= (x_{11} - x_{21})^2 + (x_{12} - x_{22})^2 = \\
 &= \left(2\sqrt{\frac{T}{2}}\right)^2 = 2T
 \end{aligned}$$

FSK jel

$M = 2$ (bináris rendszer), elemi jelek:

$$\begin{aligned}
 x_1(t) &= \cos(\omega_1 t) \\
 x_2(t) &= \cos(\omega_2 t) \quad t \in [0, T)
 \end{aligned}$$

Az ortonormált bázisfüggvények előállítására:

$$x_1(t) = a\phi_1(t),$$

$$a^2 = \int_0^T x_1^2(t) dt = \frac{T}{2},$$

ha $2\omega_1 T = \pi + l\pi$. Ebből

$$\phi_1(t) = \sqrt{\frac{2}{T}} \cos(\omega_1 t), \quad x_{11} = \sqrt{\frac{T}{2}}, \quad x_{12} = 0.$$

A Gram–Schmidt-ortogonalizálás lépéseit követve:

$$x_2(t) = b\varphi_1(t) + c\varphi_2(t),$$

$$\begin{aligned} b &= \int_0^T x_2(t)\varphi_1(t) dt = \\ &= \sqrt{\frac{2}{T}} \int_0^T \cos(\omega_2 t) \cos(\omega_1 t) dt = \\ &= \sqrt{\frac{2}{T}} \frac{1}{2} \int_0^T (\cos((\omega_1 - \omega_2)t) + \cos((\omega_1 + \omega_2)t)) dt = \\ &= \sqrt{\frac{2}{T}} \frac{1}{2} \left[\frac{\sin((\omega_1 - \omega_2)t)}{\omega_1 - \omega_2} + \frac{\sin((\omega_1 + \omega_2)t)}{\omega_1 + \omega_2} \right]_0^T \\ &= \sqrt{\frac{2}{T}} \frac{1}{2} \frac{\sin((\omega_1 - \omega_2)T)}{\omega_1 - \omega_2} \end{aligned}$$

ha $(\omega_1 + \omega_2)T = \pi + n\pi$. Bevezetve a $\Delta\omega = \omega_1 - \omega_2$ jelölést ($\omega_1 > \omega_2$)

$$b = \sqrt{\frac{2}{T}} \frac{1}{2} T \frac{\sin(\Delta\omega T)}{\Delta\omega T} = \sqrt{\frac{T}{2}} \frac{\sin(\Delta\omega T)}{\Delta\omega T}.$$

Az ortogonalizálás következő lépése c meghatározása.

$$x_2(t) - b\varphi_1(t) = c\varphi_2(t),$$

$$x_2(t) - \sqrt{\frac{T}{2}} \frac{\sin(\Delta\omega T)}{\Delta\omega T} \sqrt{\frac{2}{T}} \cos(\omega_1 t) = c\varphi_2(t).$$

Ebből

$$\begin{aligned} c^2 &= \int_0^T (x_2(t) - b\varphi_1(t))^2 dt = \\ &= \int_0^T x_2^2(t) dt + b^2 \int_0^T \varphi_1^2(t) dt - 2b \int_0^T x_2(t)\varphi_1(t) dt, \end{aligned}$$

tagonként pedig

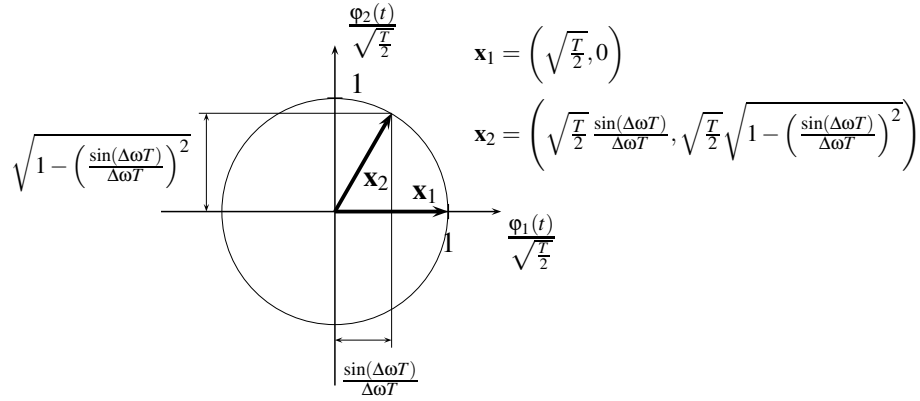
$$\int_0^T x_2^2(t) dt = \frac{T}{2},$$

ha $2\omega_2 T = \pi + h\pi$.

$$b^2 \int_0^T \varphi_1^2(t) dt = b^2,$$

és

$$\begin{aligned} 2b \int_0^T x_2(t)\varphi_1(t) dt &= 2b \int_0^T \cos(\omega_2 t) \sqrt{\frac{2}{T}} \cos(\omega_1 t) dt = \\ &= 2b \sqrt{\frac{2}{T}} \int_0^T \cos(\omega_2 t) \cos(\omega_1 t) dt = 2b \cdot b, \end{aligned}$$



1.4. ábra. Az FSK jel vektoros ábrázolása

lásd korábban. Így a c értéke könnyen meghatározható:

$$c^2 = \frac{T}{2} - b^2 = \frac{T}{2} - \frac{T}{2} \left(\frac{\sin(\Delta\omega T)}{\Delta\omega T} \right)^2,$$

vagyis

$$x_{21} = \sqrt{\frac{T}{2}} \frac{\sin(\Delta\omega T)}{\Delta\omega T}, \quad x_{22} = \sqrt{\frac{T}{2}} \sqrt{1 - \left(\frac{\sin(\Delta\omega T)}{\Delta\omega T} \right)^2}.$$

A $\varphi_2(t)$ függvény a fentiek alapján megadható, hiszen

$$\varphi_2(t) = \frac{1}{c} (x_2(t) - b\varphi_1(t)) = \frac{\cos(\omega_2 t) - \sqrt{\frac{T}{2}} \frac{\sin(\Delta\omega T)}{\Delta\omega T} \sqrt{\frac{2}{T}} \cos(\omega_1 t)}{\sqrt{\frac{T}{2}} \sqrt{1 - \left(\frac{\sin(\Delta\omega T)}{\Delta\omega T} \right)^2}}.$$

A jeleket ismét egy kétdimenziós vektortérben tudjuk ábrázolni (1.4. ábra). Ez a vektortér azonban semmilyen kapcsolatban nem áll a fázisrésszel, amely a megszokott konstellációs diagramok ábrázolására szolgál. Az ábráról leolvasható a vektorok közötti távolság:

$$\begin{aligned} \|\mathbf{x}_1 - \mathbf{x}_2\|^2 &= \|x_1(t) - x_2(t)\|^2 = \\ &= (x_{11} - x_{21})^2 + (x_{12} - x_{22})^2 = \\ &= \frac{T}{2} \left(1 - \frac{\sin(\Delta\omega T)}{\Delta\omega T} \right)^2 + \frac{T}{2} \left(1 - \left(\frac{\sin(\Delta\omega T)}{\Delta\omega T} \right)^2 \right) = \\ &= \frac{T}{2} \left(2 - 2 \frac{\sin(\Delta\omega T)}{\Delta\omega T} \right) = \\ &= T \left(1 - \frac{\sin(\Delta\omega T)}{\Delta\omega T} \right) \end{aligned}$$

Néhány speciális eset elemzése:

1. $\Delta\omega T = \pi, \Delta\omega = \frac{\pi}{T}$

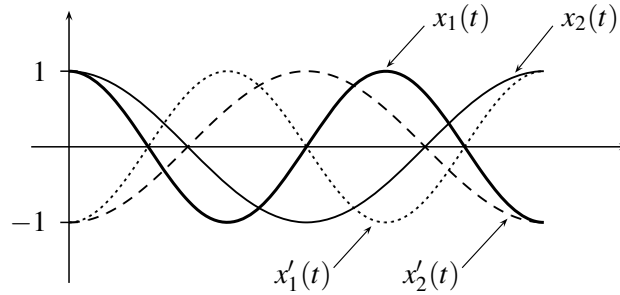
Ekkor a két jel ($x_1(t)$ és $x_2(t)$) ortogonális, pontosabban ez az a legkisebb frekvencialöket, amelyenél ez fennáll. Éppen ezért ennek a modulációnak Minimum Shift Keying (MSK) a neve.

$$\|x_1(t) - x_2(t)\|^2 = T$$

2. $\Delta\omega T = 2\pi, \Delta\omega = \frac{2\pi}{T}$

A két jel ekkor is ortogonális. A moduláció neve Fast Frequency Shift Keying (FFSK).

$$\|x_1(t) - x_2(t)\|^2 = T$$



1.5. ábra. Az MSK jel időfüggvényei

3. $\Delta\omega T = k\pi$, $k = 3, 4, \dots$, $\Delta\omega = \frac{k\pi}{T}$
A jelek ortogonálisak.

$$\|x_1(t) - x_2(t)\|^2 = T$$

4. Maximális a két jel közötti euklideszi távolság, ha

$$\frac{\sin(\Delta\omega T)}{\Delta\omega T} = \text{minimális}$$

Ennek az értéke

$$\frac{\sin(\Delta\omega T)}{\Delta\omega T} = -0.21723, \quad \Delta\omega T = 4.4934 \text{ [rad]}, \quad \Delta\omega = \frac{4.4934}{T}.$$

$$\|x_1(t) - x_2(t)\|^2 = 1.217T$$

5. $\Delta\omega T = \frac{\pi}{2}$, $\Delta\omega = \frac{\pi}{2T}$

$$\|x_1(t) - x_2(t)\|^2 = T \left(1 - \frac{2}{\pi}\right)$$

Az MSK jel további elemzése

Az MSK jel időfüggvényei az 1.5. ábrán láthatóak, ha $T = \frac{1}{1200}$, $\omega_1 = 2\pi \cdot 1800$, $\omega_2 = 2\pi \cdot 1200$, $\Delta\omega = 2\pi \cdot 600 = \frac{\pi}{T} = 1200\pi$.

A nem folytonos fázisú esetben csak $x_1(t)$ és $x_2(t)$ létezik. A jelek fázisváltozását az ún. fázisfán ábrázolhatjuk (1.6. ábra, ahol a logikai „0”-t az $x_2(t)$, míg a logikai „1”-et az $x_1(t)$ jellel visszük át). Az ábrából jól látható, hogy a logikai „1” átvitele után π értékű fázisugrás van a jelben, ami spektrálisan előnytelen. A fázisváltozást az

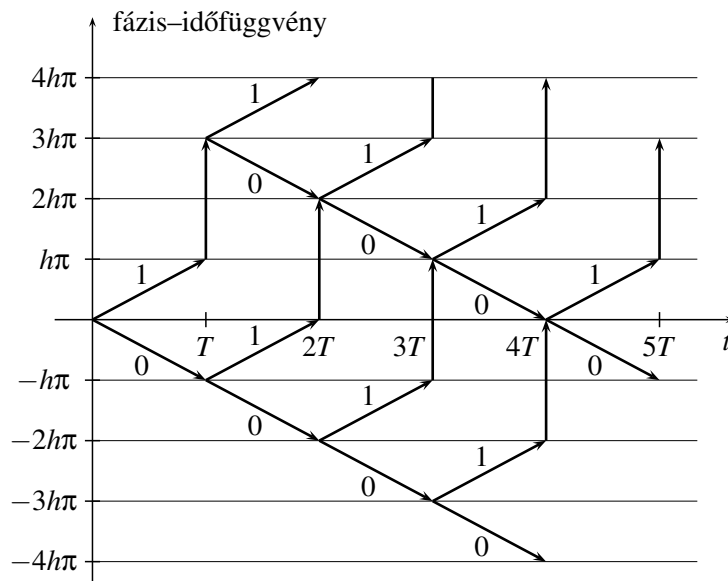
$$\omega_0 = \frac{\omega_1 + \omega_2}{2} = 2\pi \cdot 1500 \left[\frac{\text{rad}}{\text{sec}} \right]$$

elvi vivőhöz képest ábrázoltuk, azaz a T idő alatt az ω_1 frekvenciájú jel

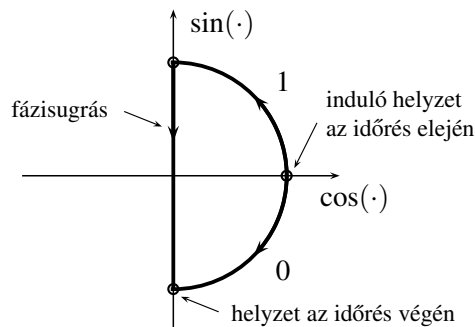
$$\Delta\varphi_1 = (\omega_1 - \omega_0)T = \frac{\omega_1 - \omega_2}{2} T = 2\pi \cdot 300 \cdot \frac{1}{1200} = \frac{\pi}{2},$$

az ω_2 frekvenciájú jel pedig

$$\Delta\varphi_2 = (\omega_2 - \omega_0)T = -\frac{\omega_1 - \omega_2}{2} T = -\frac{\pi}{2}$$



1.6. ábra. A nem folytonos fázisú MSK jel fázisfája ($h = \frac{1}{2}$)



1.7. ábra. A nem folytonos fázisú MSK jel fázor diagramja

fázistolást okoz, ami azt jelenti, hogy az ábrában $h = \frac{1}{2}$ értékű.

A jelet a szokásos fázor diagramon ábrázolva az 1.7. ábrát kapjuk.

$$\cos(\omega_1 t) = \cos(\omega_0 t + (\omega_1 - \omega_0)t) = \cos\left(\omega_0 t + \frac{\omega_1 - \omega_2}{2} t\right)$$

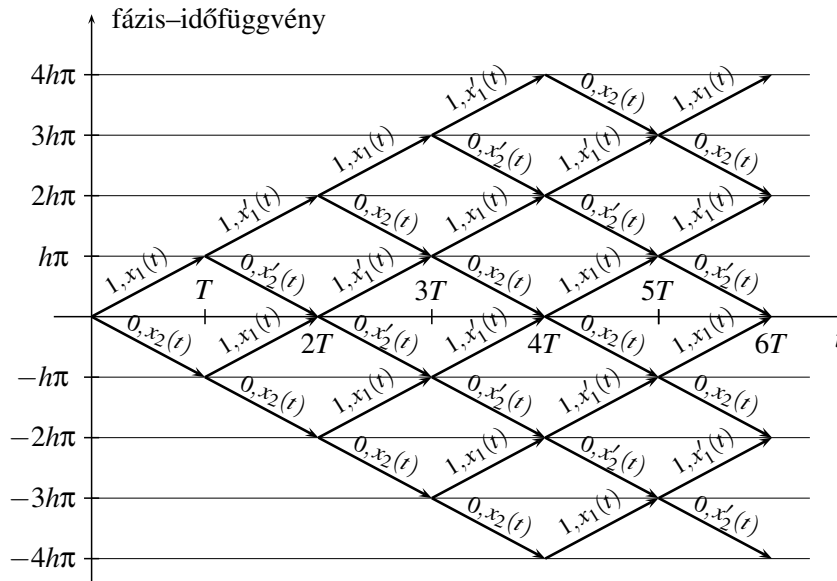
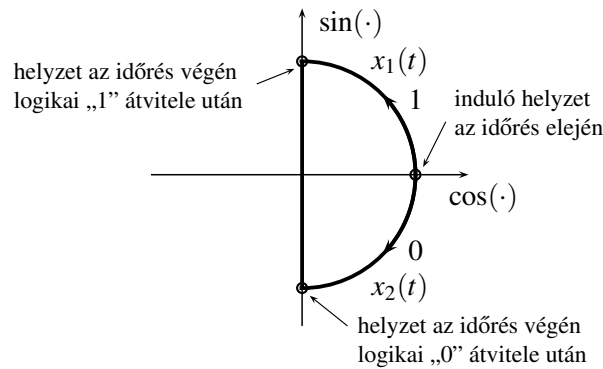
$$\cos(\omega_2 t) = \cos(\omega_0 t + (\omega_2 - \omega_0)t) = \cos\left(\omega_0 t - \frac{\omega_1 - \omega_2}{2} t\right)$$

A folytonos fázisú esetben használjuk az $x'_1(t)$ és az $x'_2(t)$ jeleket is oly módon, hogy a fázisugrások eltűnjenek. Ez annyit jelent, hogy

$$\begin{aligned} &x_1(t) \text{ után } x'_1(t) \text{ vagy } x'_2(t), \\ &x_2(t) \text{ után } x_1(t) \text{ vagy } x_2(t), \\ &x'_1(t) \text{ után } x_1(t) \text{ vagy } x_2(t), \\ &x'_2(t) \text{ után } x'_1(t) \text{ vagy } x'_2(t) \end{aligned}$$

következhethet, és a logikai „1” értéket az $x_1(t)$ vagy $x'_1(t)$, a logikai „0” értéket az $x_2(t)$ vagy $x'_2(t)$ hordozza. Ebben az esetben a fázisfa az 1.8. ábrán látható.

Az 1.9. ábrán szereplő fázor diagramból is egyértelműen eltűnik a π értékű fázisugrás.

1.8. ábra. A folytonos fázisú MSK jel fázisfája ($h = \frac{1}{2}$)

1.9. ábra. A folytonos fázisú MSK jel fázor diagramja

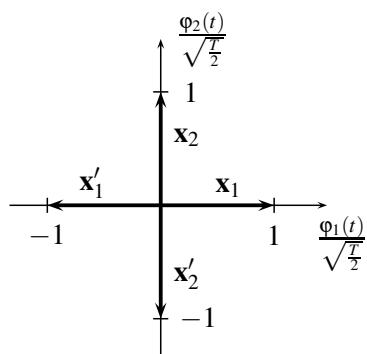
Természetesen mindkét esetre igaz az, hogy a jel burkolója állandó (constant envelop, állandó burkolójú moduláció). Ennek az a fő előnye, hogy az ilyen jel modulációs tartalma nem változik meg egy memóriamentes nemlineáris torzítás hatására, azaz a jelet nemlineáris erősítő fokozattal is lehet erősíteni. Az ilyen erősítőnek jobb a hatásfoka a lineáris erősítőkéénél.

A folytonos fázisú eset vektortérbeli jellemzéséhez két újabb elemi jelet kell ábrázolni (1.10. ábra), azaz

$$\text{„1”} \implies \mathbf{x}_1 = \left(\sqrt{\frac{T}{2}}, 0 \right) = -\mathbf{x}'_1$$

$$\text{„0”} \implies \mathbf{x}_2 = \left(0, \sqrt{\frac{T}{2}} \right) = -\mathbf{x}'_2$$

A vektortérbeli elrendezés lényegében azonos a QPSK esetével, de tudjuk, hogy itt a rendszer bináris. Ez annyit jelent, hogy a forrás logikai értékeihez az elemi jeleket nem direkt módon rendeljük hozzá, hanem ez a hozzárendelés függ a modulátor korábbi állapotaitól is. A rendszer tehát *memóriával rendelkezik*.



1.10. ábra. A folytonos fázisú MSK jel vektoros ábrázolása

2. fejezet

Az optimális koherens vevő struktúrája és hibaaránya fehér Gauss-zajos csatornában

2.1. Az optimális demodulálási szabály

Definíciók:

- $\{\pi_k\}$, $k = 1, \dots, M$ a forrás üzeneteinek a priori valószínűségei
 - $p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k)$ a vett jel feltételes valószínűségi sűrűségfüggvénye, feltéve, hogy a k -adik jelet küldték
- A feladat az, hogy a $\{\pi_k\}$ a priori valószínűségek és a $p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k)$ feltételes valószínűségi sűrűségfüggvények ismeretében egy adott \mathbf{r} vett jelvektor esetében adjunk optimális becslést a k indexre, azaz becsüljük meg azt optimálisan, hogy éppen melyik üzenetet küldték. Az optimalizálási kritérium pedig a hibavalószínűség minimalizálása, vagyis az, hogy legyen

$$P_e = \mathbf{P}(\hat{k} \neq k) = \text{minimális.}$$

Az optimális megoldást az ún. Bayes-döntés adja, amely szerint

$$\hat{k} = \underset{k}{\operatorname{argmax}} \{\pi_k p_{\mathbf{r}}(\mathbf{r} | \mathbf{x}_k)\}.$$

Ez a szabály azt mondja, hogy akkor döntünk éppen a k üzenetre, ha a vett \mathbf{r} vektor helyén pontosan a $\pi_k p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k)$ szorzat lesz a maximális. Formálisan tehát a feladatot úgy tudjuk megoldani, hogy az $\mathbf{y} = \mathbf{r}$ helyen kiszámítjuk az összes $\pi_m p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_m)$, $m = 1, \dots, M$ értéket, és ezek közül kiválasztjuk a maximálisat, majd annak indexével becsüljük az üzenet indexét.

A megoldás szemléltetésére vezessük be az ún. döntési tartomány fogalmát. Jelöljük Y -nal az \mathbf{r} vett jelvektorok terét, és \mathbf{y} -nal e tér egyik elemét. Jelölje Y_k azon \mathbf{y} vektorok halmazát, melyekre igaz, hogy $\mathbf{r} = \mathbf{y}$ esetében a fenti Bayes-döntés alkalmazásával éppen $\hat{k} = k$ döntést hozzuk, azaz a k -adik jelre döntünk. Nyilvánvaló, hogy $Y_k \cap Y_m = \emptyset$, ha $m \neq k$, és $\bigcup_{k=1}^M Y_k = Y$. A k -adik döntési tartományt tehát az

$$\begin{aligned} Y_k &= \{\mathbf{y} : \pi_k p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k) > \pi_m p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_m), \forall m \neq k\} = \\ &= \{\mathbf{y} : \ln \pi_k + \ln p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k) > \ln \pi_m + \ln p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_m), \forall m \neq k\} = \\ &= \left\{ \mathbf{y} : \ln \frac{\pi_k}{\pi_m} + \ln \frac{p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k)}{p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_m)} > 0, \forall m \neq k \right\} \end{aligned}$$

összefüggéssel lehet definiálni. (Folytonos feltételes valószínűségi sűrűségfüggvények esetén az $\{Y_k\}$ döntési tartományok határfelületei nulla mértékű halmazok, így nem kell foglalkozni velük.)

A döntési tartományokat fehér Gauss-zajos csatorna esetén a következőképpen határozhatjuk meg. A vett jelvektor feltételes valószínűségi sűrűségfüggvényeit a

$$p_r(\mathbf{y} | \mathbf{x}_k) = \left(\frac{1}{\sqrt{\pi N_0}} \right)^N \prod_{j=1}^N \exp \left(-\frac{(y_j - x_{k_j})^2}{N_0} \right), \quad k = 1, \dots, M$$

alakban adhatjuk meg, így

$$\begin{aligned} Y_k &= \left\{ \mathbf{y} : \ln \frac{\pi_k}{\pi_m} - \frac{1}{N_0} \sum_{j=1}^N (y_j - x_{k_j})^2 + \frac{1}{N_0} \sum_{j=1}^N (y_j - x_{m_j})^2 \geq 0, \forall m \neq k \right\} = \\ &= \left\{ \mathbf{y} : \ln \frac{\pi_k}{\pi_m} - \frac{1}{N_0} \sum_{j=1}^N (x_{k_j}^2 - x_{m_j}^2) + \frac{2}{N_0} \sum_{j=1}^N (x_{k_j} - x_{m_j}) y_j \geq 0, \forall m \neq k \right\}. \end{aligned}$$

Tudjuk, hogy

$$E_k = \sum_{j=1}^N x_{k_j}^2 \quad \text{és} \quad E_m = \sum_{j=1}^N x_{m_j}^2,$$

így

$$Y_k = \left\{ \mathbf{y} : \ln \frac{\pi_k}{\pi_m} - \frac{E_k - E_m}{N_0} + \frac{2}{N_0} \sum_{j=1}^N (x_{k_j} - x_{m_j}) y_j \geq 0, \forall m \neq k \right\}.$$

Megjegyezzük, hogy az Y_k definíciójában szereplő egyenlőtlenségben a határfelületről elmondottak szerint $>$ vagy \geq jelet egyaránt használhatunk.

2.2. Példák a döntési tartományok kiszámítására

QPSK típusú jelek ($N = 2, M = 4$) azonos energiákkal, de különböző a priori valószínűségekkel

Jelvektorok: $\mathbf{x}_1 = (-1, 0)$, $\mathbf{x}_2 = (0, 1)$, $\mathbf{x}_3 = (1, 0)$, $\mathbf{x}_4 = (0, -1)$.

$E_1 = E_2 = E_3 = E_4$, $\pi^{(1)} = \pi_2 = \pi_4 < \pi_1 = \pi_3 = \pi^{(2)}$

Célunk a határfelületek meghatározása. Esetünkben a döntési tartományokat az

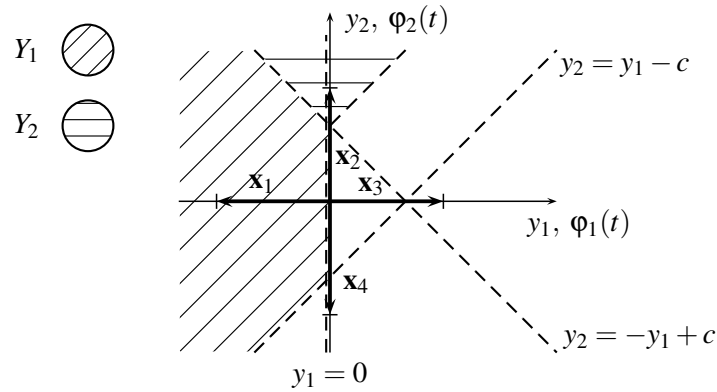
$$Y_k = \left\{ \mathbf{y} : \ln \frac{\pi_k}{\pi_m} + \frac{2}{N_0} \sum_{j=1}^N (x_{k_j} - x_{m_j}) y_j \geq 0, \forall m \neq k \right\}.$$

kifejezés határozza meg. A k -adik döntési tartomány kiszámítása végrehajtható úgy is, hogy páronként megoldjuk az

$$\ln \frac{\pi_k}{\pi_m} + \frac{2}{N_0} \sum_{j=1}^N (x_{k_j} - x_{m_j}) y_j \geq 0, \quad m \neq k$$

egyenlőtlenségeket, és azután megkeressük az Y tér azon részhalmazát, amelyre az összes egyenlőtlenség érvényes. Természetesen elegendő a határfelületeket leírni

$$\ln \frac{\pi_k}{\pi_m} + \frac{2}{N_0} \sum_{j=1}^N (x_{k_j} - x_{m_j}) y_j = 0, \quad m \neq k$$



2.1. ábra. QPSK típusú jelek döntési tartományai különböző a priori valószínűségek esetén

egyenletekkel foglalkozunk, vagy azok normált változataival

$$-\frac{N_0}{2} \ln \frac{\pi_k}{\pi_m} = \sum_{j=1}^N (x_{kj} - x_{mj})y_j, \quad m \neq k.$$

Első lépésben legyen $k = 1$ és $m = 2$.

$$-\frac{N_0}{2} \ln \frac{\pi_1}{\pi_2} \leq (x_{11} - x_{21})y_1 + (x_{12} - x_{22})y_2 = -y_1 - y_2,$$

vagyis

$$y_1 + y_2 \leq \frac{N_0}{2} \ln \frac{\pi_1}{\pi_2} > 0; \quad y_2 \leq -y_1 + c; \quad c > 0.$$

Második lépésben legyen $k = 1$ és $m = 3$.

$$0 = -\frac{N_0}{2} \ln \frac{\pi_1}{\pi_3} \leq (x_{11} - x_{31})y_1 + (x_{12} - x_{32})y_2 = -2y_1,$$

vagyis

$$0 \leq -2y_1; \quad y_1 \leq 0.$$

Harmadik lépésben legyen $k = 1$ és $m = 4$.

$$-\frac{N_0}{2} \ln \frac{\pi_1}{\pi_4} \leq (x_{11} - x_{41})y_1 + (x_{12} - x_{42})y_2 = -y_1 + y_2,$$

vagyis

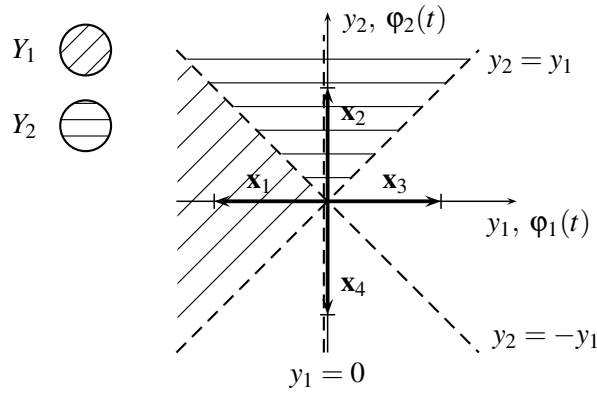
$$y_1 - y_2 \leq \frac{N_0}{2} \ln \frac{\pi_1}{\pi_4} > 0; \quad y_2 \geq y_1 - c; \quad c > 0.$$

Szimmetria miatt a többi tartományt nem kell kiszámolni. Az így megkapott döntési tartományok a 2.1. ábrán láthatók.

QPSK típusú jelek ($N = 2, M = 4$) azonos energiákkal és azonos a priori valószínűségekkel

A jelkészlet azonos az előbbi példában használttal, de $\pi_1 = \pi_2 = \pi_3 = \pi_4 = \frac{1}{4}$ és $E_1 = E_2 = E_3 = E_4$. Egyenleteink ebben az esetben még egyszerűbb alakra hozhatók, hiszen

$$Y_k = \left\{ \mathbf{y} : \sum_{j=1}^N (x_{kj} - x_{mj})y_j \geq 0, \forall m \neq k \right\},$$



2.2. ábra. QPSK típusú jelek döntési tartományai azonos a priori valószínűségek esetén

vagyis páronként csak a

$$\sum_{j=1}^N (x_{k_j} - x_{m_j}) y_j \geq 0, \quad m \neq k$$

egyenlőtlenségekkel kell foglalkoznunk.

Első lépésben legyen $k = 1$ és $m = 2$.

$$y_1 + y_2 \leq 0; \quad y_2 \leq -y_1.$$

Második lépésben legyen $k = 1$ és $m = 3$.

$$y_1 \leq 0.$$

Harmadik lépésben legyen $k = 1$ és $m = 4$.

$$y_1 - y_2 \leq 0; \quad y_2 \geq y_1.$$

Szimmetria miatt a többi tartományt nem kell kiszámolni. Az így megkapott döntési tartományok a 2.2. ábrán láthatók.

2.3. Az optimális koherens vevő felépítése

Értelmezzük a döntési szabályunkat más módon:

$$\hat{k} = \underset{k}{\operatorname{argmax}} \left\{ \left(\ln \pi_k - \frac{E_k}{N_0} \right) \frac{N_0}{2} + \sum_{j=1}^N x_{k_j} r_j \right\}.$$

Ez a felírás utasítást ad az optimális vevő felépítésére (2.3. ábra).

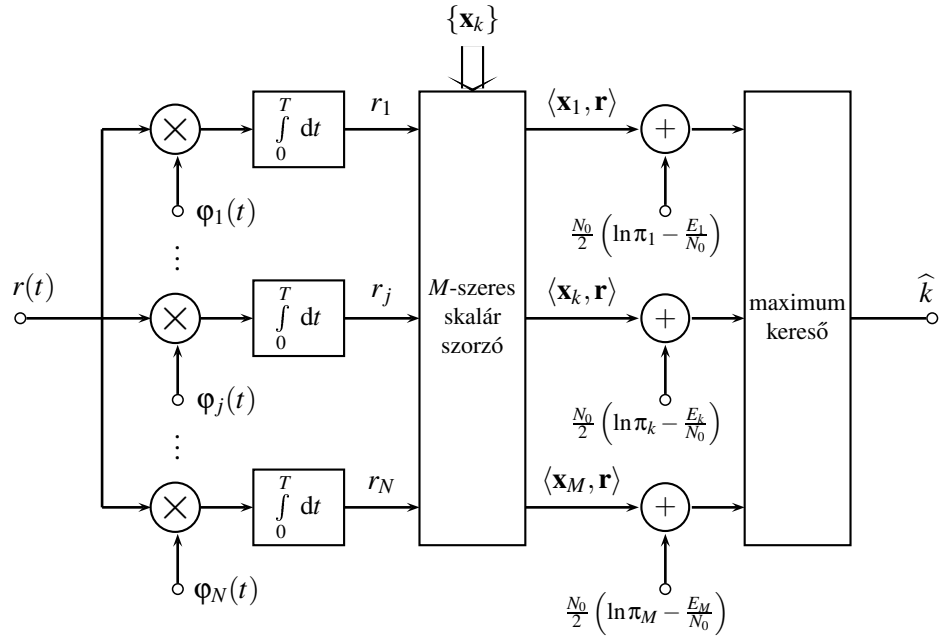
Tudjuk viszont, hogy a skalárszorzatot másképpen is elő lehet állítani, mivel

$$\sum_{j=1}^N x_{k_j} r_j = \int_0^T x_k(t) r(t) dt.$$

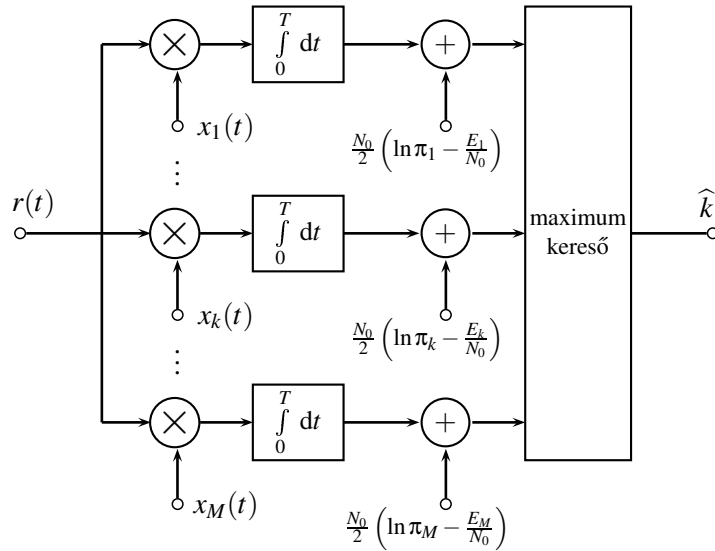
Éppen ezért a döntési szabályt le lehet írni a következőképpen is:

$$\hat{k} = \underset{k}{\operatorname{argmax}} \left\{ \left(\ln \pi_k - \frac{E_k}{N_0} \right) \frac{N_0}{2} + \int_0^T x_k(t) r(t) dt \right\},$$

amiből az optimális vevő alternatív struktúrája adódik (2.4. ábra). Ezt a vevőt korrelációs vevőnek nevezik. Jól látható, hogy nem azonos a priori valószínűségek esetén az optimális vételhez *ismerni kell a csatornában lévő zaj N_0 paraméterét.*



2.3. ábra. Az optimális koherens vevő felépítése



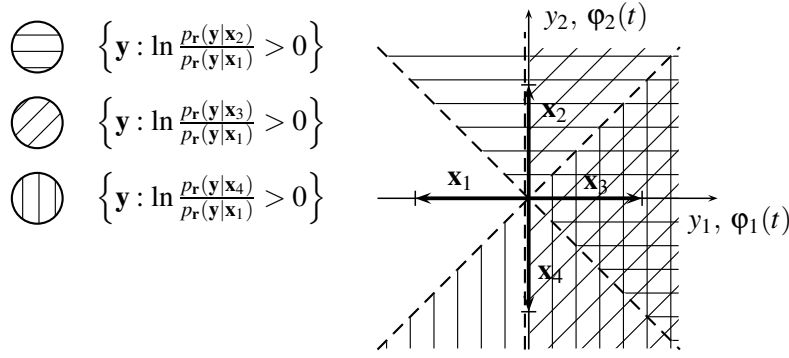
2.4. ábra. Az optimális koherens vevő alternatív felépítése

2.4. Az optimális koherens vevő hibavalószínűségeinek meghatározása

Jelöljük P_{e_k} -val a k -adik üzenet átvitelekor keletkező hibavalószínűséget, azaz legyen

$$P_{e_k} = \mathbf{P}(\mathbf{r} \in \bar{Y}_k | \mathbf{x}_k) = 1 - \mathbf{P}(\mathbf{r} \in Y_k | \mathbf{x}_k) = \int_{\mathbf{y} \in \bar{Y}_k} p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k) d\mathbf{y},$$

ahol \bar{Y}_k az Y_k döntési tartomány komplementere, $\bar{Y}_k \cup Y_k = Y$.

2.5. ábra. Az \bar{Y}_1 előállítás halmazunióval

Jelölje P_e az átlagos hibavalószínűséget, tehát

$$P_e = \sum_{k=1}^M \pi_k P_{e_k}.$$

Egyenletes a priori valószínűségeloszlás esetén

$$P_e = \frac{1}{M} \sum_{k=1}^M P_{e_k}.$$

A korábbi definíciók alapján tudjuk, hogy

$$\bar{Y}_k = \{ \mathbf{y} : \ln(p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_m)) > \ln(p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k)), \text{ legalább egy } m \neq k \text{ esetén} \},$$

vagyis

$$\bar{Y}_k = \bigcup_{\substack{m=1 \\ m \neq k}}^M \left\{ \mathbf{y} : \ln \frac{p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_m)}{p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k)} > 0 \right\}.$$

A 2.5. ábrán az \bar{Y}_1 előállítás látható.

P_{e_k} -ra felső becslést ad a

$$P_{e_k} = \mathbf{P}(\mathbf{r} \in \bar{Y}_k | \mathbf{x}_k) \leq \sum_{\substack{m=1 \\ m \neq k}}^M \mathbf{P}\left(\ln \frac{p_{\mathbf{r}}(\mathbf{r} | \mathbf{x}_m)}{p_{\mathbf{r}}(\mathbf{r} | \mathbf{x}_k)} > 0 | \mathbf{x}_k\right) = \sum_{\substack{m=1 \\ m \neq k}}^M P_e(k \rightarrow m),$$

ugyanis igaz, hogy

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B) \leq \mathbf{P}(A) + \mathbf{P}(B),$$

és definíciószerűen

$$P_e(k \rightarrow m) = \mathbf{P}(\text{csak a } k \text{ és } m \text{ üzenet közül választva } k \text{ helyett } m\text{-re döntünk} | \mathbf{x}_k).$$

Legyen $\pi_k = \frac{1}{M}$, $k = 1, \dots, M$, így

$$P_e(k \rightarrow m) = \mathbf{P}\left(\frac{2}{N_0} \sum_{j=1}^N (x_{k_j} - x_{m_j}) r_j < \frac{E_k - E_m}{N_0} | \mathbf{x}_k\right).$$

Mivel $\mathbf{r} = \mathbf{x}_k + \mathbf{v}$, ha a k -adik üzenetet küldték, akkor az \mathbf{r} vektor független, \mathbf{x}_k várható értékű és komponensenként $\frac{N_0}{2}$ szórásnégyzetű valószínűségi vektorváltozó, így

$$\eta = \frac{2}{N_0} \sum_{j=1}^N (x_{k_j} - x_{m_j})(x_{k_j} + v_j)$$

is Gauss-eloszlású valószínűségi változó μ várható értékkel és σ_η^2 szórásnégyzettel.

$$P_e(k \rightarrow m) = \mathbf{P} \left(\eta < \frac{E_k - E_m}{N_0} \right),$$

azaz a $P_e(k \rightarrow m)$ valószínűséget úgy számolhatjuk, hogy meghatározzuk, az η Gauss-eloszlású valószínűségi változó mekkora eséllyel marad az $\frac{E_k - E_m}{N_0}$ küszöb alatt.

Az η valószínűségi változó várható értékét és szórását az alábbi egyenletekből határozhatjuk meg:

$$\mu = \mathbf{E} \{ \eta \} = \mathbf{E} \left\{ \frac{2}{N_0} \sum_{j=1}^N (x_{k_j} - x_{m_j})(x_{k_j} + v_j) \right\} = \frac{2}{N_0} \sum_{j=1}^N (x_{k_j} - x_{m_j})x_{k_j}$$

$$\begin{aligned} \sigma_\eta^2 &= \mathbf{E} \{ (\eta - \mathbf{E} \{ \eta \})^2 \} = \\ &= \mathbf{E} \left\{ \left(\frac{2}{N_0} \sum_{j=1}^N (x_{k_j} - x_{m_j})v_j \right)^2 \right\} = \\ &= \frac{4}{N_0^2} \mathbf{E} \left\{ \sum_{j=1}^N \sum_{l=1}^N (x_{k_j} - x_{m_j})(x_{k_l} - x_{m_l})v_j v_l \right\} = \\ &= \frac{4}{N_0^2} \sum_{j=1}^N \sum_{l=1}^N (x_{k_j} - x_{m_j})(x_{k_l} - x_{m_l}) \mathbf{E} \{ v_j v_l \} = \\ &= \frac{4}{N_0^2} \frac{N_0}{2} \sum_{j=1}^N (x_{k_j} - x_{m_j})^2 = \\ &= \frac{2}{N_0} \sum_{j=1}^N (x_{k_j} - x_{m_j})^2 \end{aligned}$$

$$P_e(k \rightarrow m) = \mathbf{P} \left(\eta < \frac{E_k - E_m}{N_0} \right) = \int_{-\infty}^{\frac{E_k - E_m}{N_0}} \frac{1}{\sqrt{2\pi}\sigma_\eta} \exp \left(-\frac{(z - \mu)^2}{2\sigma_\eta^2} \right) dz$$

Használjuk fel az alábbi definíciót:

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{z^2}{2} \right) dz = \int_{-\infty}^{-x} \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{z^2}{2} \right) dz,$$

vagy

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-z^2) dz, \quad \operatorname{erfc}(x) = 2Q(\sqrt{2}x).$$

Behelyettesítve az $x = \frac{z - \mu}{\sigma_\eta}$, $dz = \sigma_\eta dx$ új változót

$$P_e(k \rightarrow m) = \int_{-\infty}^{\frac{1}{\sigma_\eta} \left(\frac{E_k - E_m}{N_0} - \mu \right)} \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{x^2}{2} \right) dx = Q \left(\left(\frac{E_m - E_k}{N_0} + \mu \right) \frac{1}{\sigma_\eta} \right).$$

A Q függvény argumentumának elemzése után

$$\begin{aligned}
 \frac{1}{\sigma_\eta} \left(\frac{E_m - E_k}{N_0} + \mu \right) &= \frac{1}{N_0} \frac{\sum_{j=1}^N (x_{m_j}^2 - x_{k_j}^2) + 2 \sum_{j=1}^N (x_{k_j} - x_{m_j}) x_{k_j}}{\sqrt{\frac{2}{N_0} \sum_{j=1}^N (x_{k_j} - x_{m_j})^2}} = \\
 &= \frac{\frac{1}{N_0} \sum_{j=1}^N (x_{k_j} - x_{m_j})^2}{\sqrt{\frac{2}{N_0} \sum_{j=1}^N (x_{k_j} - x_{m_j})^2}} = \\
 &= \frac{1}{\sqrt{2N_0}} \sqrt{\sum_{j=1}^N (x_{k_j} - x_{m_j})^2} = \\
 &= \frac{\|\mathbf{x}_k - \mathbf{x}_m\|}{\sqrt{2N_0}},
 \end{aligned}$$

vagyis

$$P_e(k \rightarrow m) = Q \left(\frac{\|\mathbf{x}_k - \mathbf{x}_m\|}{\sqrt{2N_0}} \right),$$

azaz a páronkénti hibaválószerűség csak a jelvektorok euklideszi távolságától és a fehér Gauss-zaj spektrális sűrűségétől függ.

2.5. Példák a hibaválószerűség számítására

BPSK jel

$$N = 1, M = 2, \pi_k = \frac{1}{2}$$

Elemi jelek:

$$\begin{aligned}
 x_1(t) &= \sqrt{2P} \cos(\omega_0 t), \\
 x_2(t) &= -\sqrt{2P} \cos(\omega_0 t),
 \end{aligned}$$

ahol P a teljesítmény és $t \in [0, T)$.

$$\varphi(t) = \sqrt{\frac{2}{T}} \cos(\omega_0 t), \quad 2\omega_0 T = \pi + n\pi.$$

$$\begin{aligned}
 \|\mathbf{x}_1 - \mathbf{x}_2\|^2 &= \int_0^T (x_1(t) - x_2(t))^2 dt = \\
 &= 2P \int_0^T (2 \cos(\omega_0 t))^2 dt = \\
 &= 2P \int_0^T 4 \cdot \frac{1 + \cos(2\omega_0 t)}{2} dt = \\
 &= 2P \cdot 2T = 4PT = 4E,
 \end{aligned}$$

így

$$P_e = P_e(1 \rightarrow 2) = Q\left(\frac{\sqrt{4E}}{\sqrt{2N_0}}\right) = Q\left(\sqrt{2}\sqrt{\frac{E}{N_0}}\right).$$

Bináris esetben a számítás pontos, mivel csak páronkénti összehasonlítást kell végezni.

QPSK jel

$$N = 2, M = 4, \pi_k = \frac{1}{4}$$

Elemi jelek:

$$\begin{aligned} x_1(t) &= \sqrt{2P} \cos(\omega_0 t) = -x_2(t), \\ x_3(t) &= \sqrt{2P} \sin(\omega_0 t) = -x_4(t), \end{aligned}$$

ahol P a teljesítmény és $t \in [0, T)$.

$$\begin{aligned} \varphi_1(t) &= \sqrt{\frac{2}{T}} \cos(\omega_0 t), \quad 2\omega_0 T = \pi + n\pi \\ \varphi_2(t) &= \sqrt{\frac{2}{T}} \sin(\omega_0 t). \end{aligned}$$

A páronkénti hibavalószínűségek számítása:

$$P_e(1 \rightarrow m) = Q\left(\frac{\|\mathbf{x}_1 - \mathbf{x}_m\|}{\sqrt{2N_0}}\right), \quad m \neq 1,$$

$$\|\mathbf{x}_1 - \mathbf{x}_2\|^2 = \int_0^T (x_1(t) - x_2(t))^2 dt = 4PT = 4E,$$

$$\|\mathbf{x}_1 - \mathbf{x}_3\|^2 = \|\mathbf{x}_1 - \mathbf{x}_4\|^2 = \int_0^T (x_1(t) - x_3(t))^2 dt = 2P\left(\frac{T}{2} + \frac{T}{2}\right) = 2PT = 2E.$$

$$P_e = P_{e_1} = P_{e_2} = P_{e_3} = P_{e_4} \leq 2Q\left(\sqrt{\frac{E}{N_0}}\right) + Q\left(\sqrt{2}\sqrt{\frac{E}{N_0}}\right) < 3Q\left(\sqrt{\frac{E}{N_0}}\right)$$

A BPSK és QPSK jelek hibaarányainak összehasonlítása esetén fontos megjegyezni, hogy bináris forrás esetén, ha T_b a bináris forrás szimbólumideje, akkor

- BPSK esetén $T = T_b$,
- QPSK esetén viszont $T = 2T_b$.

Ennek az a következménye, hogy azonos P jelteljesítmények mellett

- BPSK esetén $E = PT = PT_b = E_b$,
- QPSK esetén $E = PT = P(2T_b) = 2E_b$,

ezért a hibaarányokra igaz, hogy

- BPSK esetén $P_e = Q\left(\sqrt{2}\sqrt{\frac{E_b}{N_0}}\right)$, míg
- QPSK esetén $P_e < 3Q\left(\sqrt{2}\sqrt{\frac{E_b}{N_0}}\right)$,

azaz a két rendszer hibaaránya közel azonos.

MSK jel

Nem folytonos fázisú, $N = 2$, $M = 2$, $\pi_k = \frac{1}{2}$

Elemi jelek:

$$\begin{aligned} x_1(t) &= \sqrt{2P} \cos(\omega_1 t), \\ x_2(t) &= \sqrt{2P} \cos(\omega_2 t), \end{aligned}$$

ahol P a teljesítmény és $t \in [0, T)$.

$$\Delta\omega = \omega_1 - \omega_2 > 0, \quad \Delta\omega T = \pi$$

$$\begin{aligned} \varphi_1(t) &= \sqrt{\frac{2}{T}} \cos(\omega_1 t), & 2\omega_1 T &= \pi + l\pi, & (\omega_1 + \omega_2)T &= \pi + n\pi. \\ \varphi_2(t) &= \sqrt{\frac{2}{T}} \cos(\omega_2 t), & 2\omega_2 T &= \pi + h\pi, \end{aligned}$$

$$\begin{aligned} \|\mathbf{x}_1 - \mathbf{x}_2\|^2 &= \int_0^T (x_1(t) - x_2(t))^2 dt = \\ &= 2P \left(\frac{T}{2} + \frac{T}{2} \right) = 2PT = 2E, \end{aligned}$$

$$P_e = Q\left(\frac{\|\mathbf{x}_1 - \mathbf{x}_2\|}{\sqrt{2N_0}}\right) = Q\left(\frac{\sqrt{2E}}{\sqrt{2N_0}}\right) = Q\left(\sqrt{\frac{E}{N_0}}\right).$$

2.6. A hibaarány közelítő számítása

Korábban láttuk, hogy

$$P_{e_k} \leq \sum_{\substack{m=1 \\ m \neq k}}^M P_e(k \rightarrow m) = \sum_{\substack{m=1 \\ m \neq k}}^M Q\left(\frac{\|\mathbf{x}_k - \mathbf{x}_m\|}{\sqrt{2N_0}}\right) \leq (M-1) Q\left(\frac{\|\mathbf{x}_k - \mathbf{x}_m\|_{\min}}{\sqrt{2N_0}}\right).$$

A $Q(x)$ függvény viselkedését bármely x esetében az alábbi egyenlőtlenséggel lehet leírni:

$$\left(1 - \frac{1}{x^2}\right) \frac{\exp\left(-\frac{x^2}{2}\right)}{\sqrt{2\pi x}} < Q(x) < \frac{\exp\left(-\frac{x^2}{2}\right)}{\sqrt{2\pi x}}, \quad x > 0,$$

amelyből nagy x -ek esetére a

$$Q(x) \simeq \frac{\exp\left(-\frac{x^2}{2}\right)}{\sqrt{2\pi x}}$$

közelítés adódik, mivel az alsó és felső korlát is aszimptotikusan ehhez az értékhez tart. Ez azt jelenti, hogy nagy $\frac{\|\mathbf{x}_k - \mathbf{x}_m\|}{\sqrt{2N_0}}$ esetén a

$$Q\left(\frac{\|\mathbf{x}_k - \mathbf{x}_m\|}{\sqrt{2N_0}}\right) \simeq \frac{1}{\sqrt{2\pi} \frac{\|\mathbf{x}_k - \mathbf{x}_m\|}{\sqrt{2N_0}}} \exp\left(-\frac{1}{2} \frac{\|\mathbf{x}_k - \mathbf{x}_m\|^2}{2N_0}\right)$$

közelítés érvényes, vagyis az $\frac{\|\mathbf{x}_k - \mathbf{x}_m\|}{\sqrt{2N_0}}$ növekedésével a függvény értéke igen nagy sebességgel csökken. A $Q(x)$ függvény további tulajdonságait a fejezet végén foglaljuk össze.

Tekintsük példaképpen a BPSK hibaarányát! Ekkor

$$P_e = Q\left(\sqrt{2} \sqrt{\frac{E}{N_0}}\right) \simeq \frac{1}{\sqrt{2\pi} \sqrt{2} \sqrt{\frac{E}{N_0}}} \exp\left(-\frac{E}{N_0}\right).$$

Ha most E értékét E' -re változtatjuk, akkor

$$P'_e = Q\left(\sqrt{2}\sqrt{\frac{E'}{N_0}}\right),$$

és a két hibaarány hányadosa

$$\frac{P'_e}{P_e} = \frac{Q\left(\sqrt{2}\sqrt{\frac{E'}{N_0}}\right)}{Q\left(\sqrt{2}\sqrt{\frac{E}{N_0}}\right)} \simeq \sqrt{\frac{E}{E'}} \exp\left(-\frac{E' - E}{N_0}\right)$$

értékű lesz. Ha például $\frac{E}{N_0} = 10$ és $\frac{E'}{N_0} = 20$, akkor

$$\frac{P'_e}{P_e} \simeq \sqrt{\frac{1}{2}} \exp(-10) \simeq \sqrt{\frac{1}{2}} 10^{-4.4}.$$

Éppen emiatt azt lehet mondani, hogy kis hibaarányú (nagy jel–zaj viszonyú) rendszerekben a hibaarányt a jelek (jelvektorok) közötti minimális euklideszi távolság határozza meg, ugyanis a többi additív tag elhanyagolható.

A korábbi felső becslést alkalmazva, az átlagos hibaarányt a

$$\begin{aligned} P_e &= \sum_{k=1}^M \pi_k P_{e_k} \leq \\ &\leq \sum_{k=1}^M \pi_k \sum_{\substack{m=1 \\ m \neq k}}^M P_e(k \rightarrow m) = \\ &= \sum_{k=1}^M \pi_k \sum_{\substack{m=1 \\ m \neq k}}^M Q\left(\frac{\|\mathbf{x}_k - \mathbf{x}_m\|}{\sqrt{2N_0}}\right) \leq \\ &\leq \sum_{k=1}^M \pi_k (M-1) Q\left(\frac{\|\mathbf{x}_k - \mathbf{x}_m\|_{\min}}{\sqrt{2N_0}}\right) = \\ &= (M-1) Q\left(\frac{\|\mathbf{x}_k - \mathbf{x}_m\|_{\min}}{\sqrt{2N_0}}\right) \end{aligned}$$

kifejezéssel lehet számolni.

A $Q(x)$ függvény előbb említett tulajdonságait figyelembe véve gyakran alkalmazzák az alábbi közelítést:

$$\begin{aligned} P_e &= \sum_{k=1}^M \pi_k P_{e_k} \leq \\ &\leq \sum_{k=1}^M \pi_k \sum_{\substack{m=1 \\ m \neq k}}^M P_e(k \rightarrow m) = \\ &= \sum_{k=1}^M \pi_k \sum_{\substack{m=1 \\ m \neq k}}^M Q\left(\frac{\|\mathbf{x}_k - \mathbf{x}_m\|}{\sqrt{2N_0}}\right) \approx \\ &\approx \sum_{k=1}^M \pi_k a_{k \min} Q\left(\frac{\|\mathbf{x}_k - \mathbf{x}_m\|_{\min}}{\sqrt{2N_0}}\right) = \\ &= K_{\min} Q\left(\frac{\|\mathbf{x}_k - \mathbf{x}_m\|_{\min}}{\sqrt{2N_0}}\right), \end{aligned}$$

ahol $a_{k\min}$ adott k esetén az $\|\mathbf{x}_k - \mathbf{x}_m\|_{\min}$ előfordulásának a számossága $m = 1, \dots, M$, $m \neq k$ esetén, és $K_{\min} = \mathbf{E}\{a_{k\min}\} = \sum_{k=1}^M \pi_k a_{k\min}$.

Visszatérve a QPSK korábbi példájára, $k = 1$ esetén

$$\|\mathbf{x}_1 - \mathbf{x}_m\| = \begin{cases} 2\sqrt{E}, & m = 2 \\ \sqrt{2E}, & m = 3 \\ \sqrt{2E}, & m = 4 \end{cases}$$

$a_{1\min} = 2$, szimmetria okok miatt természetesen $a_{\min} = 2$, és $\|\mathbf{x}_k - \mathbf{x}_m\|_{\min} = \sqrt{2E}$. Éppen ezért a közelítő hibaarány kifejezése most

$$P_e \approx 2Q\left(\sqrt{\frac{E}{N_0}}\right)$$

a pontos felső korlát helyett, amelyet a

$$P_e \leq 2Q\left(\sqrt{\frac{E}{N_0}}\right) + Q\left(\sqrt{2}\sqrt{\frac{E}{N_0}}\right)$$

kifejezéssel számítottunk. Ez azért volt megtehető, mert nagy jel-zaj viszony esetén a második tag az elsőhöz képest elhanyagolható (lásd a korábbi példát).

A $Q(x)$ függvény további tulajdonságai

A Gauss hibavalószínűség függvény ($Q(x)$ függvény) definíció szerint a

$$Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz$$

kifejezéssel adható meg.

A függvény fő tulajdonsága, hogy nagy argumentumok esetén az értéke a

$$Q(x) \approx \exp\left(-\frac{x^2}{2}\right)$$

kifejezéssel közelíthető.

A következőkben bemutatjuk a függvény speciális közelítéseit:

- 1. Közelítés

$$Q(x) < \frac{1}{\sqrt{2\pi}x} \exp\left(-\frac{x^2}{2}\right),$$

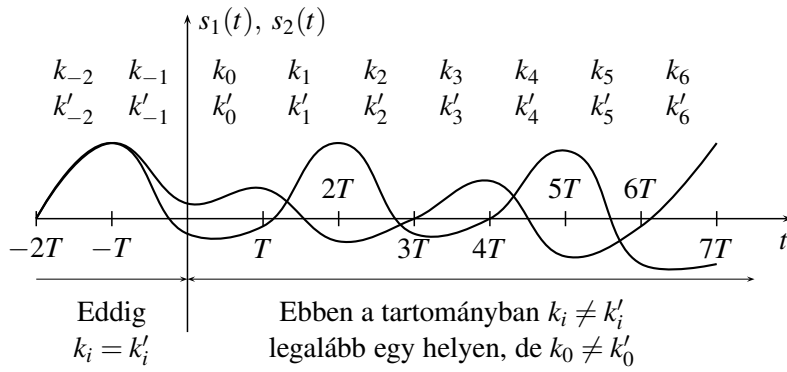
- 2. Közelítés

$$Q(x) > \left(1 - \frac{1}{x^2}\right) \frac{1}{\sqrt{2\pi}x} \exp\left(-\frac{x^2}{2}\right),$$

- 3. Közelítés

$$Q(x) \leq \frac{1}{4} \exp\left(-\frac{x^2}{2}\right),$$

ha x elegendően nagy.



2.6. ábra. A minimális euklideszi távolság számítása

2.7. A koherens modulációs rendszerek általános jellemzése

A rendszerek eddigi leírásánál nem kötöttük ki azt, hogy a forrás ξ_{k_i} szimbólumai milyen komplexitásúak, és arra sem adtunk előírást, hogy a T szimbólumidő milyen hosszúságú. Bármikor megtehetjük azt, hogy több, például z darab szimbólumot összefogunk, és ezekhez a zT időrésben elemi jeleket rendelünk. Nyilvánvaló, hogy ebben az esetben összesen M^z számú különböző üzenetet kell átvinnünk a csatornán, azaz ennyi elemi jelet kell generálnunk. Természetesen z minden határon túl növelhető, így a teljes modulált jel különböző realizációit is elemi jelnek tekinthetjük, amelyeket a $\{\xi_{k_i}\}$ sorozatok egy-egy realizációjához rendeltünk. A jelek vektortérbeli leírása most is alkalmazható, és értelmezni lehet általános esetben is az $\|\mathbf{x}_k - \mathbf{x}_m\|_{\min}$ ún. minimális euklideszi távolságot, amely a rendszer hibarányára ad iránymutatást.

Ezzel a megközelítéssel a rendszer globális minőségi jellemzését azokra az esetekre is általánosíthatjuk, amikor a modulációs eljárás bonyolultabb (pl. memória van a modulátorban, lásd folytonos fázisú MSK). Célunk ilyen esetekben a minimális euklideszi távolság meghatározása, és segítségével a hibarány becslése.

Jelöljük $s(t, \{\xi_{k_i}\})$ és $s(t, \{\xi_{k'_i}\})$ -vel a modulált jel két realizációját, melyeket rendre az egymástól legalább egy szimbólumban különböző $\{\xi_{k_i}\}$ és $\{\xi_{k'_i}\}$ szimbólumsorozatokhoz rendeltünk, és jelölje d_{\min} a minimális euklideszi távolságot.

A d_{\min} számítása elvégezhető, mindössze végre kell hajtani az alábbi minimalizálást:

$$d_{\min}^2 = \min_{\{k_i\}, \{k'_i\}} \int_{-\infty}^{\infty} \left(s(t, \{\xi_{k_i}\}) - s(t, \{\xi_{k'_i}\}) \right)^2 dt,$$

ahol $\{k_i\}$ és $\{k'_i\}$ nem azonos, $i = -\infty, \dots, -1, 0, 1, \dots, \infty$, $k_i, k'_i \in \{1, \dots, M\}$.

A feladat praktikus megoldását a 2.6. ábrán szemléltetjük az

$$\begin{aligned} s_1(t) &= s(t, \{\xi_{k_i}\}) \\ s_2(t) &= s(t, \{\xi_{k'_i}\}) \end{aligned}$$

jelölések bevezetése után.

A feladat tehát úgy is megfogalmazható, hogy válasszuk meg a $\{k_i\}$ és $\{k'_i\}$ sorozatokat oly módon, hogy

$$\begin{aligned} k_i &= k'_i, & i < 0 \\ k_i &\neq k'_i, & i = 0 \\ k_i &= k'_i \text{ vagy } k_i \neq k'_i, & i > 0. \end{aligned}$$

Rendeljük ezekhez a sorozatokhoz a modulált jelet, és keressük meg a lehetséges modulált jelek közötti minimális euklideszi távolságot.

Egyszerűen belátható, hogy abban az esetben, ha becsülni akarjuk a rendszerre jellemző hibarányt, akkor a korábbiak szerint ez a

$$P_e \approx K_{\min} Q \left(\frac{d_{\min}}{\sqrt{2N_0}} \right)$$

kifejezés segítségével megtehető.

2.8. Példák a koherens modulációs rendszerek általános jellemzésére

Modulációs rendszerek általános elemi jelekkel

Paraméterek:

- ξ_{k_i} -hez az $x_{k_i}(t - iT)$ elemi jelet rendeljük, ahol $i = -\infty, \dots, -1, 0, 1, \dots, \infty$, $k_i \in \{1, \dots, M\}$.
- Az átviendő üzenetek lehetséges értékkészlete M méretű.
- A szimbólumidő T .
- A modulátor memóriamentes.

A modulált jel általános alakja:

$$s(t, \{\xi_{k_i}\}) = \sum_{i=-\infty}^{\infty} x_{k_i}(t - iT),$$

$x_k(t)$ tartója $[0, T)$.

A minimális euklideszi távolság meghatározásához szükséges jelek az alábbiak:

$$\begin{aligned} s_1(t) &= \sum_{i=-\infty}^{\infty} x_{k_i}(t - iT) \\ s_2(t) &= \sum_{i=-\infty}^{\infty} x_{k'_i}(t - iT), \end{aligned}$$

és teljesül, hogy

$$\begin{aligned} k_i &= k'_i, & i < 0 \\ k_i &\neq k'_i, & i = 0 \\ k_i &= k'_i, & i > 0, \end{aligned}$$

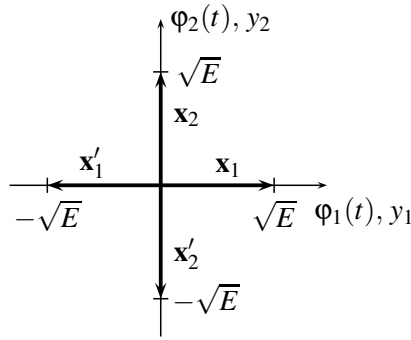
mivel nyilvánvaló, hogy a

$$d_{\min}^2 = \min_{\{k_i\}, \{k'_i\}} \int_{-\infty}^{\infty} (s_1(t) - s_2(t))^2 dt$$

kifejezést a $\{k_i\}$ és $\{k'_i\}$ sorozat éppen ilyen választás mellett minimalizálja, azaz

$$\begin{aligned} d_{\min}^2 &= \min_{\substack{k_0, k'_0 \\ k_0 \neq k'_0}} \int_{-\infty}^{\infty} (x_{k_0}(t) - x_{k'_0}(t))^2 dt = \\ &= \min_{\substack{k, l \\ k \neq l}} \int_{-\infty}^{\infty} (x_k(t) - x_l(t))^2 dt = \\ &= \min_{\substack{k, l \\ k \neq l}} \|\mathbf{x}_k - \mathbf{x}_l\|^2. \end{aligned}$$

Mindez azt jelenti, hogy a minimális euklideszi távolságot az időrekenkénti elemi jelek között euklideszi távolságuk közül a minimális kiválasztásával lehet meghatározni. Hasonló módon kaphatjuk meg az a_{\min} értékét is (lásd a korábbi példát).



2.7. ábra. Folytonos fázisú MSK jelek vektortérbeli ábrázolása

Folytonos fázisú MSK moduláció

$M = 4$, bináris átviteli rendszer

Paraméterek: $\xi_{k_i} \in [0, 1]$, $\Delta\omega T = \pi$, $\Delta\omega = \omega_1 - \omega_2 > 0$.

Elemi jelek:

$$\begin{aligned} x_1(t) &= \sqrt{2P} \cos(\omega_1 t), \\ x_1'(t) &= -\sqrt{2P} \cos(\omega_1 t), \\ x_2(t) &= \sqrt{2P} \cos(\omega_2 t), \\ x_2'(t) &= -\sqrt{2P} \cos(\omega_2 t), \end{aligned}$$

ahol P a teljesítmény és $t \in [0, T)$.

Az elemi jelvektorok rendszere ($E = PT$):

$$\begin{aligned} \mathbf{x}_1 &= (\sqrt{E}, 0) \\ \mathbf{x}_1' &= (-\sqrt{E}, 0) \\ \mathbf{x}_2 &= (0, \sqrt{E}) \\ \mathbf{x}_2' &= (0, -\sqrt{E}) \end{aligned}$$

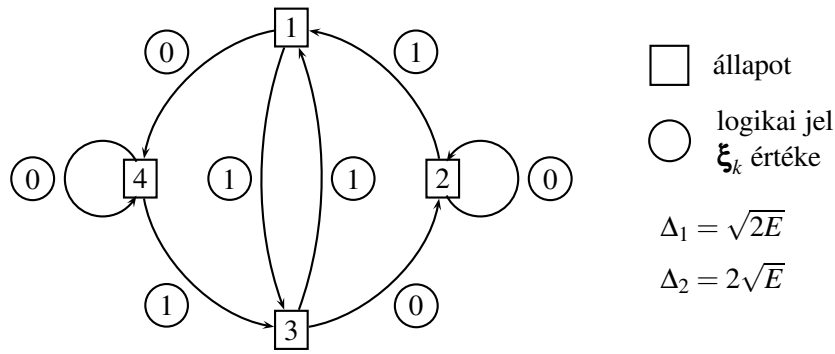
A jelek vektortérbeli ábrázolása a 2.7. ábrán látható.

A jelek hozzárendelése a forrásszimbólumokhoz attól függ, hogy az előző időrésben melyik jelet adtuk le. A logikai „1” értéket az \mathbf{x}_1 vagy \mathbf{x}_1' , a logikai „0” értéket az \mathbf{x}_2 vagy \mathbf{x}_2' hordozza. Hozzárendelési szabályok:

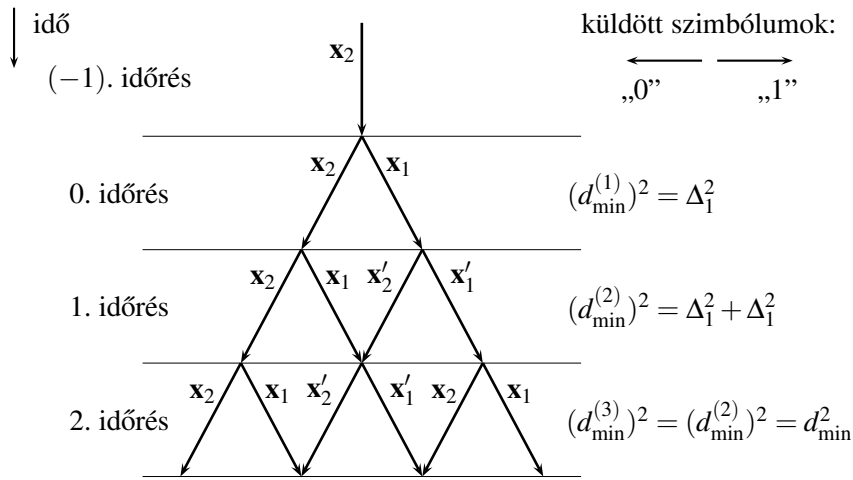
$$\begin{aligned} x_1(t) \text{ után csak } x_1'(t) \text{ vagy } x_2'(t), \\ x_2(t) \text{ után csak } x_1(t) \text{ vagy } x_2(t), \\ x_1'(t) \text{ után csak } x_1(t) \text{ vagy } x_2(t), \\ x_2'(t) \text{ után csak } x_1'(t) \text{ vagy } x_2'(t) \text{ következhet.} \end{aligned}$$

A modulátor állapotérbeli leírása. A rendszerben a különböző elemi jelek csak bizonyos sorrendben követhetik egymást, így lehetőség van arra, hogy a minimális euklideszi távolság a korábban elemzett nem folytonos fázisú, két elemi jelet ($x_1(t)$ és $x_2(t)$) használó MSK rendszeréhez képest növekedjen. Mint korábban láttuk, $d_{\min}^2 = 2E$ volt ez az érték, ami nem más, mint az \mathbf{x}_1 és az \mathbf{x}_2 távolsága.

A rendszerben a modulátor állapotát a korábbi időrésben küldött jel határozza meg. Ennek alapján



2.8. ábra. Az MSK modulátor állapotátmeneti diagramja



2.9. ábra. Távolságszámítás MSK moduláció esetén

felrajzolható a rendszer állapotátmeneti diagramja (2.8. ábra). Az állapotok definíciója:

- 1 állapot: x_1 volt az előző jel
- 2 állapot: x_2 volt az előző jel
- 3 állapot: x'_1 volt az előző jel
- 4 állapot: x'_2 volt az előző jel

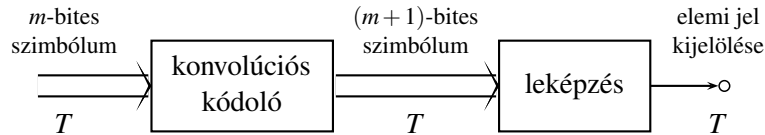
Az ábra alapján világosan látszik, hogy ha például az összehasonlításához kiválasztott $s_1(t)$ és $s_2(t)$ jelek elágazása a nulladik időrésben a 2 állapotból indul, akkor annak a két jelnek a távolságát kell elemezni, amely a nulladik időrésben x_2 illetve x_1 elemi jelekkel kezdődik.

A vizsgálatot a 2.9. ábra diagramja segíti. Az első időrésben elő kell állítani az

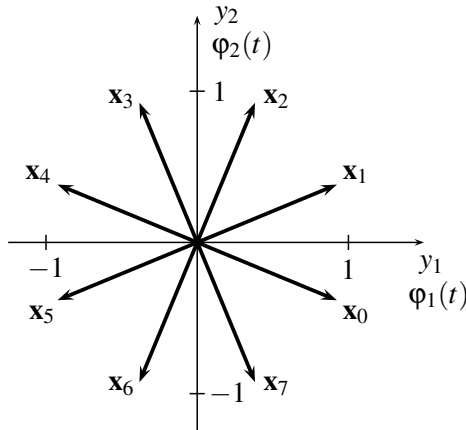
$$\begin{aligned} \|x_2 - x'_2\| &= \|x_1 - x'_1\| = \Delta_2 \\ \|x_1 - x'_2\| &= \|x'_1 - x_2\| = \Delta_1 \end{aligned}$$

távolságokat. Egyszerűen belátható, hogy az elágazás után a

$$\begin{array}{cc} \begin{matrix} 0 & 0 & 0 \\ x_2 & x_2 & x_2 \end{matrix} & \begin{matrix} 0 & 1 & 0 \\ x_2 & x_1 & x'_2 \end{matrix} \\ & \text{vagy} \\ \begin{matrix} 1 & 1 & 0 \\ x_1 & x'_1 & x_2 \end{matrix} & \begin{matrix} 1 & 0 & 0 \\ x_1 & x'_2 & x'_2 \end{matrix} \end{array}$$



2.10. ábra. Az Ungerboeck-kód előállítás (m = 2)



2.11. ábra. Az Ungerboeck-kód jelkészlete

sorozatokhoz tartozik a minimális euklideszi távolság, amely $d_{\min}^2 = 4E$.

Hasonló eredményre jutunk akkor is, ha a (-1) -edik időrésben az x_1 jel küldését feltételezzük. Fontos megjegyezni, hogy az eljárással, azaz a memória elhelyezésével a modulátorban, az eredeti nem folytonos fázisú MSK jelhez képest a minimális euklideszi távolságot $\sqrt{2}$ -szeresre, a jel effektív energiáját kétszeresre növeltük. Ez azt eredményezi, hogy kisebb jelteljesítménnyel lehet azonos hibaarányal kommunikálni.

A rendszerhez a következők kellene:

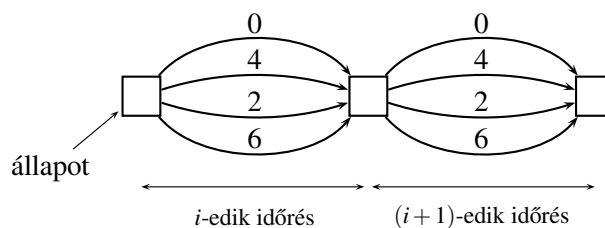
- memória és szekvenciális logika a modulátorban (véges állapotgép a jelek sorrendjének meghatározására);
- az átvinni kívánt szimbólumok számosságánál bővebb jeltér.

Az Ungerboeck-kód 4PSK esetben

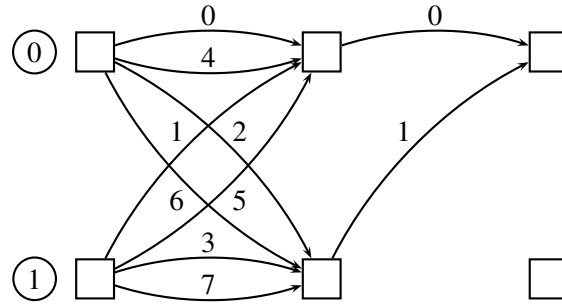
$M = 8$, az átvitt szimbólumok számossága 4

Cél az adatsebesség megtartása mellett a minimális euklideszi távolság növelése az eredeti QPSK rendszerhez képest.

A modulátor általános felépítése a 2.10. ábrán látható. 8PSK-nak megfelelő a jelkészlet és kétdimenziós a jeltér (2.11. ábra).



2.12. ábra. Az egy belső állapotú rendszer távolságszámítása



2.13. ábra. A két belső állapotú rendszer távolságszámítása

Az elemi jelek távolság alapú csoportosítása:

$$\begin{aligned}\Delta_0 &= \|\mathbf{x}_0 - \mathbf{x}_1\| = 2 \sin \frac{\pi}{8} = 0.765, \\ \Delta_1 &= \|\mathbf{x}_0 - \mathbf{x}_2\| = \sqrt{2} = 1.414, \\ \Delta_2 &= \|\mathbf{x}_0 - \mathbf{x}_4\| = 2.\end{aligned}$$

A jelcsoportokat úgy alakítjuk ki, hogy a tagok között a távolság homogén legyen.

$$\begin{array}{ccc} \underbrace{\mathbf{x}_0, \mathbf{x}_4, \mathbf{x}_2, \mathbf{x}_6}_{C_0} & \underbrace{\mathbf{x}_1, \mathbf{x}_5, \mathbf{x}_3, \mathbf{x}_7}_{C_1} & \Delta_1 \\ & & \Delta_2 \end{array}$$

A jelek sorrendjének meghatározását befolyásoló irányelvek:

- Azonos a priori valószínűségek, szimmetrikus és reguláris jeltér (ez teljesül a 8PSK jeltér esetén).
- Egy belső állapotból induló átmenetek vagy csak a B_0 -ból, vagy csak a B_1 -ből választhatók.
- Egy belső állapotba befutó átmenetek vagy csak a B_0 -ból, vagy csak a B_1 -ből választhatók.
- Ha két állapot között vannak párhuzamos átmenetek, azokat vagy csak a C_0 -ból, vagy csak a C_1 -ből, vagy csak a C_2 -ből, vagy csak a C_3 -ből lehet választani.

Ezek az elvek biztosítják azt, hogy az egyes jelek között az euklideszi távolság minél nagyobb legyen.

Néhány eset vizsgálata. A hagyományos 4PSK, vagy egy állapotú rendszer a 2.12. ábrán látható. Itt

$$d_{\min} = \Delta_1 = 1.414.$$

A két belső állapotú rendszert a 2.13. ábrán tüntettük fel.

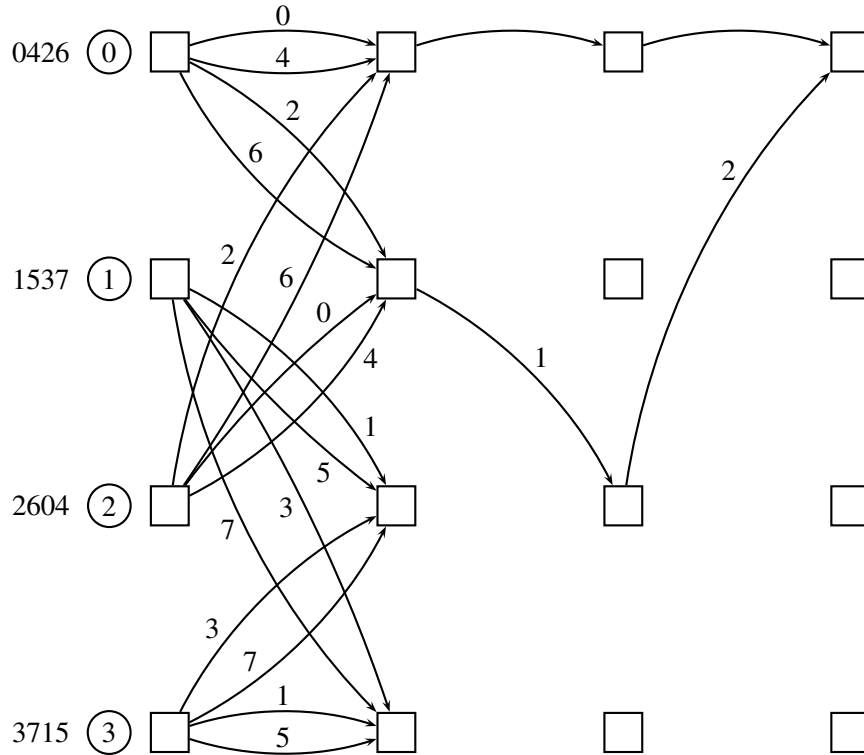
$$d_{\min} = \sqrt{\Delta_1^2 + \Delta_0^2} = 1.608,$$

amely 1.1 dB-es nyereséget jelent a 4PSK-hoz képest. Ez a távolság például az $(\mathbf{x}_0, \mathbf{x}_0)$ és az $(\mathbf{x}_2, \mathbf{x}_1)$ jelpárok távolságából adódik.

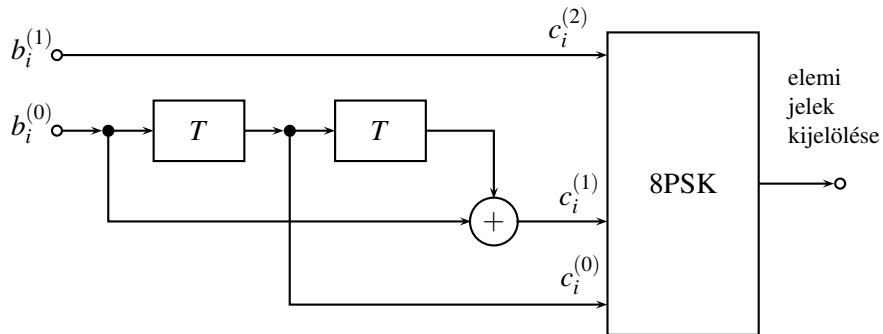
A négy belső állapotú rendszer a 2.14. ábrán látható.

$$d_{\min} = 2,$$

amely 3 dB-es javulást jelent a 4PSK-hoz képest. Ez a távolság például az $(\mathbf{x}_2, \mathbf{x}_1, \mathbf{x}_2)$ és az $(\mathbf{x}_6, \mathbf{x}_1, \mathbf{x}_2)$, vagy az $(\mathbf{x}_0, \mathbf{x}_0, \mathbf{x}_0)$ és az $(\mathbf{x}_4, \mathbf{x}_0, \mathbf{x}_0)$ sorozatok között. A feladatot a 2.15. ábra konvolúciós kódolóval oldhatjuk meg. A konvolúciós kódoló a $(b_i^{(0)}, b_i^{(1)})$ bináris szimbólum pároshoz a $(c_i^{(0)}, c_i^{(1)}, c_i^{(2)})$ bináris szimbólum hármast rendel, amely kijelöli a 8PSK jeltér megfelelő elemét. A rendszer állapotátmeneti diagramját a 2.16. ábrán adtuk meg. Az ábra jelölései:



2.14. ábra. A négy belső állapotú rendszer távolságszámítása



2.15. ábra. Négy belső állapotú konvolúciós kódoló

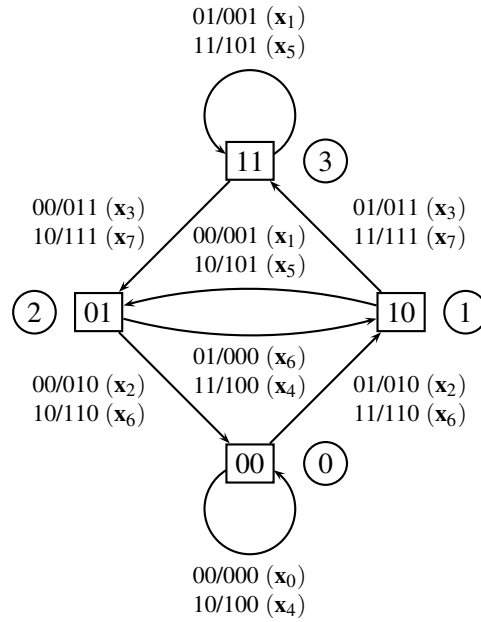
- állapot: $b_{i-1}^{(0)}, b_{i-2}^{(0)}$
- átmenet: $b_i^{(1)} b_i^{(0)} / c_i^{(2)} c_i^{(1)} c_i^{(0)} (\mathbf{x}_k)$
- állapotok sorszáma: ①

Az ábrából világosan látszik, hogy a konvolúciós kódoló éppen az állapotátmeneti diagramot valósítja meg.

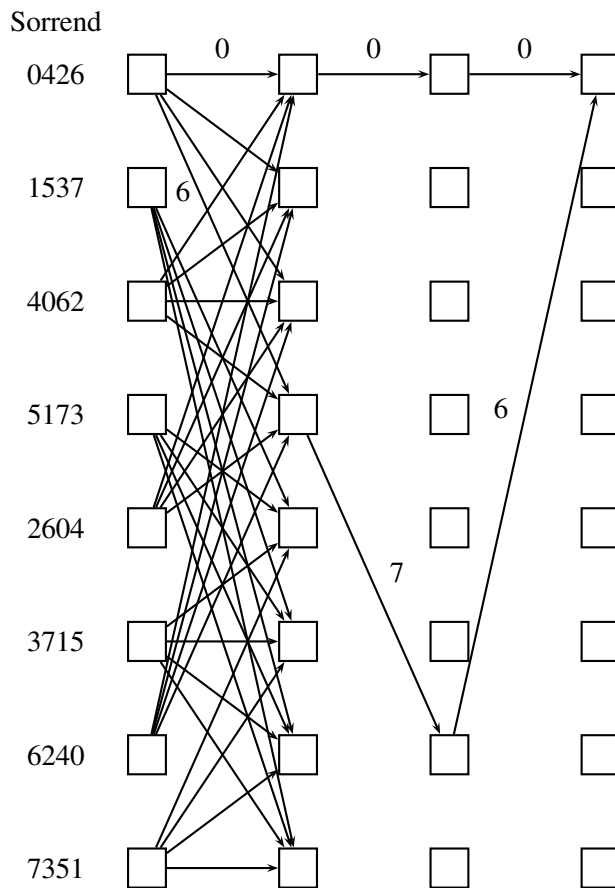
A nyolc belső állapotú rendszer a 2.17. ábrán látható. A modulátor memóriájának növelése a minimális euklideszi távolság további növelését teszi lehetővé, ugyanis

$$d_{\min} = \sqrt{\Delta_1^2 + \Delta_1^2 + \Delta_0^2} = 2.141,$$

amely 3.6 dB-es javulást jelent a 4PSK-hoz képest ($a_{\min} = 2$). Ez a távolság például az $(\mathbf{x}_0, \mathbf{x}_0, \mathbf{x}_0)$ és az $(\mathbf{x}_6, \mathbf{x}_7, \mathbf{x}_6)$ sorozatok között.



2.16. ábra. A négy belső állapotú konvolúciós kódoló állapotátmeneti diagramja



2.17. ábra. A nyolc belső állapotú rendszer távolságszámítása

3. fejezet

Az optimális nem koherens vevő struktúrája és hibaaránya

3.1. A nem koherens jelek jeltérbeli ábrázolása

Legyen a forrás által előállított szimbólumok sorozat $\{\xi_{k_i}\}$, melynek minden eleme egy M értékészletű ábécéből veszi fel az értékeit. Rendeljük a ξ_{k_i} szimbólumhoz az $x_{k_i}(t - iT)$ elemi jelet úgy, hogy $x_k(t)$ tartója most is a $[0, T)$ tartomány. Ekkor elegendő például az $i = 0$ esetet, azaz a nulladik időrést vizsgálni.

Válasszuk az elemi jeleket

$$x_k(t) = \sqrt{2} z_k(t) \cos(\omega_0 t), \quad k = 1, \dots, M$$

alakúra. Állítsuk elő Gram–Schmidt-ortogonalizálással a $\{\varphi_j(t)\}$ -t, a $\{z_k(t)\}$ függvények által generált ortonormált teljes bázist, $j = 1, \dots, N$, $N \leq M$, azaz

$$z_k(t) = \sum_{j=1}^N z_{kj} \varphi_j(t), \quad z_{kj} = \int_0^T z_k(t) \varphi_j(t) dt,$$

$$\langle \varphi_j(t), \varphi_l(t) \rangle = \begin{cases} 1, & j = l \\ 0, & j \neq l \end{cases}$$

Nem koherens rendszerben a vevőbe érkező jel vivőjének fázisáról nincs információnk. Ezért az

$$r(t) = \sqrt{2} z_k(t) \cos(\omega_0 t - \Theta) + v(t)$$

vevőbe érkező jelről feltételezzük, hogy az Θ véletlen fázissal érkezik, és Θ a $[0, 2\pi)$ tartományban egyenletes eloszlású valószínűségi változó, amely független a forrás szimbólumaitól.

A vevőbe érkező modulált jelek leírásához vezessünk be egy új $2N$ dimenziós ortonormált bázis-rendszert:

$$\begin{aligned} \Psi_{c_j}(t) &= \sqrt{2} \varphi_j(t) \cos(\omega_0 t), & j &= 1, \dots, N, \\ \Psi_{s_j}(t) &= \sqrt{2} \varphi_j(t) \sin(\omega_0 t), & j &= 1, \dots, N, \end{aligned}$$

amelyről belátható, hogy

$$\langle \Psi_{c_j}(t), \Psi_{c_l}(t) \rangle = \begin{cases} 1, & j = l \\ 0, & j \neq l \end{cases}$$

$$\langle \Psi_{s_j}(t), \Psi_{s_l}(t) \rangle = \begin{cases} 1, & j = l \\ 0, & j \neq l \end{cases}$$

és $\langle \Psi_{c_j}(t), \Psi_{s_l}(t) \rangle = 0$ minden j, l párra, ha $\varphi_j(t)$ ω_0 -hoz képest erősen sávhatárolt.

Ily módon az adó által küldött $\{x_k(t)\}$ elemi jelek a $\{\Psi_{c_j}(t)\}$ ortonormált bázis segítségével egyszerűen leírhatók, mivel

$$\begin{aligned} x_k(t) &= \sqrt{2}z_k(t) \cos(\omega_0 t) = \\ &= \sum_{j=1}^N z_{k_j} \sqrt{2} \varphi_j(t) \cos(\omega_0 t) = \\ &= \sum_{j=1}^N z_{k_j} \Psi_{c_j}(t) = \\ &= \sum_{j=1}^N x_{k_j} \Psi_{c_j}(t), \end{aligned}$$

amiből jól látható, hogy a $\{\varphi_j(t)\}$ ortonormált bázis segítségével előállított \mathbf{z}_k vektor és a $\{\Psi_{c_j}(t)\}$ ortonormált bázis segítségével előállított \mathbf{x}_k vektor minden komponense azonos. Igaz, hogy az előbbit az alapsávi ortonormált bázisfüggvényekkel ($\{\varphi_j(t)\}$), az utóbbit pedig az ún. koszinuszos irányú vivősávi ortonormált bázisfüggvényekkel ($\{\Psi_{c_j}(t)\}$) generáltuk. Mindebből látható, hogy

$$z_{k_j} = x_{k_j}$$

minden $j \in \{1, \dots, N\}$ értékre.

A koherens vételhez hasonlóan a további vizsgálatokhoz elő kell állítani a vevő bemenetére érkező $r(t)$ függvényt elégséges statisztikával leíró \mathbf{r} vektort, ami azt jelenti, hogy olyan ortonormált bázisfüggvényrendszert kell használni, amellyel az $r(t)$ -ben lévő hasznos jelet pontosan le lehet írni. A k -adik üzenet küldése esetén az $r(t)$ -ben lévő hasznos jel a

$$\sqrt{2}z_k(t) \cos(\omega_0 t - \Theta) = \sqrt{2}z_k(t) \cos(\Theta) \cos(\omega_0 t) + \sqrt{2}z_k(t) \sin(\Theta) \sin(\omega_0 t)$$

alakban adható meg, és mivel

$$z_k(t) = \sum_{j=1}^N z_{k_j} \varphi_j(t),$$

a fenti összefüggés a

$$\begin{aligned} \sqrt{2}z_k(t) \cos(\omega_0 t - \Theta) &= \\ &= \cos(\Theta) \sum_{j=1}^N z_{k_j} \sqrt{2} \varphi_j(t) \cos(\omega_0 t) + \sin(\Theta) \sum_{j=1}^N z_{k_j} \sqrt{2} \varphi_j(t) \sin(\omega_0 t) = \\ &= \cos(\Theta) \sum_{j=1}^N z_{k_j} \Psi_{c_j}(t) + \sin(\Theta) \sum_{j=1}^N z_{k_j} \Psi_{s_j}(t) \end{aligned}$$

alakban írható fel.

A kifejezésekből nyilvánvaló, hogy a vevőbe érkező hasznos jelet csak egy $2N$ -dimenziós ortonormált bázisfüggvényrendszerrel lehet teljesen leírni, ami azt jelenti, hogy a vivő fázisának ismerete nélkül a bejövő jel mindkét kvadratúra összetevőjét meg kell figyelni ahhoz, hogy a hasznos jelről minden lehetséges információhoz hozzájussunk a vevőben. A fázis ismeretének hiánya azt eredményezi, hogy a fehér Gauss-zajból kétszer annyi teljesítmény jut a vevőbe, mint koherens esetben. Az \mathbf{r} vektor egy $2N$ -dimenziós, minden irányban független, nem nulla várható értékű, $\frac{N_0}{2}$ szórásnégyzetű Gauss-eloszlású valószínűségi vektorváltozó, azaz, ha a k -adik üzenetet küldték, akkor

$$\begin{aligned} \mathbf{r} &= (\mathbf{r}_c, \mathbf{r}_s) = \\ &= (r_{c_1}, \dots, r_{c_N}, r_{s_1}, \dots, r_{s_N}) = \\ &= (\cos(\Theta)x_{k_1} + \mathbf{v}_{c_1}, \dots, \cos(\Theta)x_{k_N} + \mathbf{v}_{c_N}, \sin(\Theta)x_{k_1} + \mathbf{v}_{s_1}, \dots, \sin(\Theta)x_{k_N} + \mathbf{v}_{s_N}), \end{aligned}$$

ahol $\mathbf{v}_c = (v_{c_1}, \dots, v_{c_N})$ és $\mathbf{v}_s = (v_{s_1}, \dots, v_{s_N})$ N -dimenziós független, nulla várható értékű és $\frac{N_0}{2}$ szórásnégyzetű Gauss-eloszlású vektorváltozók.

Ennek alapján, ha Θ egy rögzített értéket vesz fel, az \mathbf{r} vektor feltételes valószínűségi sűrűségfüggvénye a

$$\begin{aligned} p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k, \Theta) &= p_{\mathbf{r}}(\mathbf{y}_c, \mathbf{y}_s | \mathbf{x}_k, \Theta) = \\ &= \left(\frac{1}{\sqrt{\pi N_0}} \right)^{2N} \prod_{j=1}^N \exp \left(-\frac{(y_{c_j} - \cos(\Theta)x_{k_j})^2 + (y_{s_j} - \sin(\Theta)x_{k_j})^2}{N_0} \right) \end{aligned}$$

alakban írható fel, amit a továbbiak előkészítése érdekében a következőképpen alakíthatunk át:

$$\begin{aligned} p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k, \Theta) &= \left(\frac{1}{\pi N_0} \right)^N \prod_{j=1}^N \exp \left(-\frac{y_{c_j}^2 + y_{s_j}^2}{N_0} \right) \exp \left(-\frac{x_{k_j}^2}{N_0} \right) \exp \left(2 \frac{\cos(\Theta)y_{c_j} + \sin(\Theta)y_{s_j}}{N_0} x_{k_j} \right) = \\ &= \left(\frac{1}{\pi N_0} \right)^N \exp \left(-\frac{\sum_{j=1}^N (y_{c_j}^2 + y_{s_j}^2)}{N_0} \right) \exp \left(-\frac{\sum_{j=1}^N x_{k_j}^2}{N_0} \right) \exp \left(2 \frac{\cos(\Theta) \sum_{j=1}^N y_{c_j} x_{k_j} + \sin(\Theta) \sum_{j=1}^N y_{s_j} x_{k_j}}{N_0} \right) = \\ &= \left(\frac{1}{\pi N_0} \right)^N \exp \left(-\frac{\|\mathbf{y}\|^2}{N_0} \right) \exp \left(-\frac{E_k}{N_0} \right) \exp \left(2 \frac{\cos(\Theta) \langle \mathbf{y}_c, \mathbf{x}_k \rangle + \sin(\Theta) \langle \mathbf{y}_s, \mathbf{x}_k \rangle}{N_0} \right). \end{aligned}$$

A korábbi vizsgálatokból tudjuk, hogy az optimális vételhez, azaz a Bayes-döntéshez a $p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k, \Theta)$ helyett a $p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k)$ valószínűségi sűrűségfüggvényre van szükségünk, amelyet a $p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k, \Theta)$ -ből egyszerűen származtathatunk, mivel általában

$$p_X(x) = \int_{a \in \mathcal{A}} p_X(x | a) p_A(a) da,$$

ahol A a feltételben szereplő valószínűségi változó, $p_A(a)$ ennek a sűrűségfüggvénye, \mathcal{A} pedig az értékkészlete. Ez alapján

$$p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k) = \int_0^{2\pi} p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k, \vartheta) p_{\Theta}(\vartheta) d\vartheta = \frac{1}{2\pi} \int_0^{2\pi} p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k, \vartheta) d\vartheta.$$

Az integrálás végrehajtásához használjuk fel azt az összefüggést, hogy

$$\frac{1}{2\pi} \int_0^{2\pi} \exp(c \cos(\vartheta)) d\vartheta = \frac{1}{2\pi} \int_0^{2\pi} \exp(c \cos(\vartheta - \Theta_0)) d\vartheta = I_0(c),$$

ahol $I_0(c)$ a másodfajú, nulladrendű Bessel-függvény a c helyen. Az $I_0(c)$ függvény tulajdonságait az alábbiakkal lehet jellemezni:

- $I_0(0) = 1$;
- $I_0(x) = 1 + \frac{x^2}{4} + \frac{x^4}{64} + \frac{x^6}{2304} + \dots$, $|x| \ll 1$;
- $I_0(x)$ $x > 0$ esetén monoton növekvő és
- $I_0(x) \simeq \left(1 + \frac{1}{8x}\right) \frac{\exp(x)}{\sqrt{2\pi x}}$, $|x| \gg 1$.

Ezek alapján a kijelölt

$$\frac{1}{2\pi} \int_0^{2\pi} \exp \left(2 \frac{\cos(\vartheta) \langle \mathbf{y}_c, \mathbf{x}_k \rangle + \sin(\vartheta) \langle \mathbf{y}_s, \mathbf{x}_k \rangle}{N_0} \right) d\vartheta$$

integrál könnyen kiszámítható. Vezessük be az

$$a = 2 \frac{\langle \mathbf{y}_c, \mathbf{x}_k \rangle}{N_0} \quad b = 2 \frac{\langle \mathbf{y}_s, \mathbf{x}_k \rangle}{N_0}$$

jelöléseket, és használjuk fel az

$$a \cos(\vartheta) + b \sin(\vartheta) = \sqrt{a^2 + b^2} \cos(\vartheta - \Theta_0)$$

összefüggést, ahol

$$\Theta_0 = \operatorname{arctg} \left(\frac{b}{a} \right).$$

Ebben az esetben

$$\frac{1}{2\pi} \int_0^{2\pi} \exp(a \cos(\vartheta) + b \sin(\vartheta)) d\vartheta = I_0 \left(\sqrt{a^2 + b^2} \right),$$

tehát

$$p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k) = \left(\frac{1}{\pi N_0} \right)^N \exp \left(-\frac{\|\mathbf{y}\|^2}{N_0} \right) \exp \left(-\frac{E_k}{N_0} \right) I_0 \left(2 \frac{\rho_k}{N_0} \right),$$

ahol

$$\rho_k = \sqrt{\langle \mathbf{y}_c, \mathbf{x}_k \rangle^2 + \langle \mathbf{y}_s, \mathbf{x}_k \rangle^2},$$

és tudjuk a korábbi vizsgálatokból, hogy

$$\begin{aligned} \langle \mathbf{y}_c, \mathbf{x}_k \rangle &= \langle \mathbf{y}_c, \mathbf{z}_k \rangle = \sum_{j=1}^N y_{c_j} x_{k_j}, \\ \langle \mathbf{y}_s, \mathbf{x}_k \rangle &= \langle \mathbf{y}_s, \mathbf{z}_k \rangle = \sum_{j=1}^N y_{s_j} x_{k_j}, \\ E_k &= \sum_{j=1}^N x_{k_j}^2. \end{aligned}$$

Az optimális vevőben a hibaarányt minimalizáló Bayes-döntés szerint

$$\hat{k} = \operatorname{argmax}_k \{ \pi_k p_{\mathbf{r}}(\mathbf{r} | \mathbf{x}_k) \},$$

ami azt jelenti, hogy akkor döntünk éppen a k üzenetre, ha

$$\pi_k \exp \left(-\frac{E_k}{N_0} \right) I_0 \left(2 \frac{\rho_k}{N_0} \right) > \pi_m \exp \left(-\frac{E_m}{N_0} \right) I_0 \left(2 \frac{\rho_m}{N_0} \right) \quad \forall m \neq k,$$

ahol

$$\rho_k = \sqrt{\langle \mathbf{r}_c, \mathbf{x}_k \rangle^2 + \langle \mathbf{r}_s, \mathbf{x}_k \rangle^2}.$$

A vevőstruktúrák elemzése előtt célszerű értelmezni az aktuális skalárszorzatokat. Tudjuk, hogy az \mathbf{r} vektor két részből áll (ezek lényegében a vevőbe érkező jelvektor kvadratúra összetevői). $\mathbf{r} = (\mathbf{r}_c, \mathbf{r}_s)$, ahol

$$\begin{aligned} \mathbf{r}_c &= (r_{c_1}, \dots, r_{c_N}), \\ \mathbf{r}_s &= (r_{s_1}, \dots, r_{s_N}). \end{aligned}$$

Maga az $r(t)$ jel (a fehér Gauss-zaj téren kívüli, elhanyagolható komponenseit nem tekintve) az

$$r(t) = r_c(t) + r_s(t) = \sum_{j=1}^N r_{c_j} \Psi_{c_j}(t) + \sum_{j=1}^N r_{s_j} \Psi_{s_j}(t)$$

formában adható meg. Tudjuk, hogy

$$x(t) = \sum_{j=1}^N x_{k_j} \Psi_{c_j}(t),$$

ezért

$$\begin{aligned} \langle \mathbf{r}_c, \mathbf{x}_k \rangle &= \int_0^T r_c(t) x_k(t) dt = \\ &= \int_0^T \left(\sum_{j=1}^N r_{c_j} \Psi_{c_j}(t) \right) \left(\sum_{l=1}^N x_{k_l} \Psi_{c_l}(t) \right) dt = \\ &= \int_0^T \left(\sum_{j=1}^N r_{c_j} \Psi_{c_j}(t) + r_{s_j} \Psi_{s_j}(t) \right) \left(\sum_{l=1}^N x_{k_l} \Psi_{c_l}(t) \right) dt = \\ &= \int_0^T r(t) x_k(t) dt = \\ &= \int_0^T r(t) \sqrt{2} z_k(t) \cos(\omega_0 t) dt, \end{aligned}$$

és hasonlóan

$$\begin{aligned} \langle \mathbf{r}_s, \mathbf{x}_k \rangle &= \int_0^T r_s(t) x_k(t) dt = \\ &= \int_0^T \left(\sum_{j=1}^N r_{s_j} \Psi_{s_j}(t) \right) \left(\sum_{l=1}^N x_{k_l} \Psi_{s_l}(t) \right) dt = \\ &= \int_0^T \left(\sum_{j=1}^N r_{c_j} \Psi_{c_j}(t) + r_{s_j} \Psi_{s_j}(t) \right) \left(\sum_{l=1}^N x_{k_l} \Psi_{s_l}(t) \right) dt = \\ &= \int_0^T r(t) \sqrt{2} z_k(t) \sin(\omega_0 t) dt, \end{aligned}$$

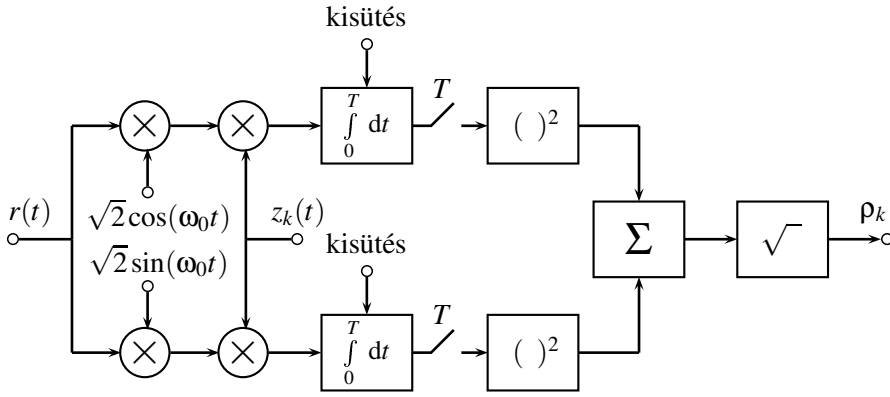
mivel

$$\sum_{l=1}^N x_{k_l} \Psi_{s_l}(t) = \sum_{l=1}^N z_{k_l} \sqrt{2} \varphi_j(t) \sin(\omega_0 t) = \sqrt{2} z_k(t) \sin(\omega_0 t).$$

3.2. Optimális nem koherens vevőstruktúrák

A fenti kifejezések alapján az optimális vevőstruktúra könnyen meghatározható, hiszen a legfontosabb feladat a ρ_k paraméterek előállítása. A 3.1. ábrán az ún. alapsávi kvadratúra korrelátoros vevő felépítést adtuk meg. Az optimális vételhez egy-egy ilyen kvadratúra korrelátor előállítja a ρ_k paramétereket, amelyeket az optimális vételhez szükséges egyenlőtlenségekbe kell helyettesíteni, és az így előállított

$$\pi_k \exp\left(-\frac{E_k}{N_0}\right) I_0\left(2 \frac{\rho_k}{N_0}\right), \quad k = 1, \dots, M$$



3.1. ábra. Az alapsávi kvadratúra korrelátoros vevő felépítése

szám M -es alapján a maximális értékű indexére kell döntenünk, azaz

$$\hat{k} = \operatorname{argmax}_k \left\{ \pi_k \exp \left(-\frac{E_k}{N_0} \right) I_0 \left(2 \frac{\rho_k}{N_0} \right) \right\}.$$

Jól látható, hogy $\pi_k = \frac{1}{M}$ esetén csak akkor tudunk optimálisan dönteni, ha ismerjük N_0 értékét (azonos a priori valószínűségek, de különböző jelenergiák). $E_k = E$ esetén is csak akkor tudunk dönteni, ha ismerjük N_0 értékét (azonos jelenergiák, de különböző a priori valószínűségek). Viszont $E_k = E$ és $\pi_k = \frac{1}{M}$ esetén (azonos jelenergiák és azonos a priori valószínűségek) elegendő a

$$\hat{k} = \operatorname{argmax}_k \left\{ I_0 \left(2 \frac{\rho_k}{N_0} \right) \right\} = \operatorname{argmax}_k \{ \rho_k \}$$

értékét kiszámítani, mivel az $I_0(x)$ függvény az $x > 0$ tartományban monoton növekvő.

Megjegyzendő, hogy a nem koherens átviteli rendszerek többségére ezek a feltételek teljesülnek (pl. OFSK, Orthogonal Frequency Shift Keying).

A kvadratúra korrelátoros vevőnek elő lehet állítani a különböző ekvivalens változatait, mivel a korrelátor egy ún. illesztett szűrővel ekvivalens módon helyettesíthető. Ennek az az alapja, hogy a ρ_k értéket csak egy adott pillanatban, a $[0, T)$ időtartomány végén lehet értelmezni, hiszen az integrálás ebben a pillanatban fejeződik be.

Hasonlítsuk össze egy alapsávi integrátor és egy alapsávi szűrő működését, ha a bejövő jel $v_c(t)$. Ekkor

$$\int_0^T v_c(\tau) z_k(\tau) d\tau$$

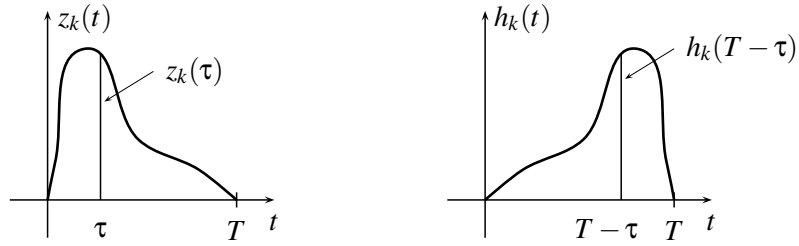
a korrelátor kimenete a T időpontban,

$$\int_{-\infty}^t v_c(\tau) h_k(t - \tau) d\tau$$

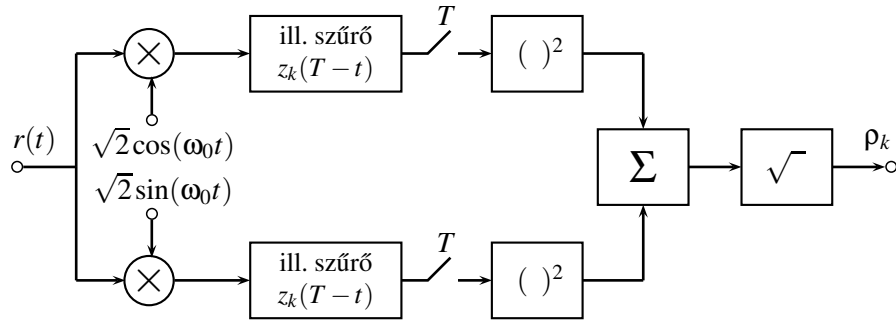
pedig a $h_k(t - \tau)$ súlyfüggvényű szűrő kimenő jele a t időpontban. A két jel a $t = T$ időpontban akkor azonos, ha

$$h_k(T - t) = \begin{cases} z_k(t), & t \in [0, T) \\ 0 & \text{egyébként} \end{cases}$$

Ezt illusztráltuk a 3.2. ábrán. Az ilyen szűrőt illesztett szűrőnek nevezzük. Megállapítható, hogy a vizsgált korrelátor és az illesztett szűrő ekvivalensek egymással.



3.2. ábra. Az alapsávi integrátor és szűrő kimenetei



3.3. ábra. Alapsávi nem koherens illesztett szűrős vevő

Megjegyezzük, hogy a korrelátoros és illesztett szűrős vevők olyan esetben is hasonlóan értelmezhetők, ha az elemi jelek tartója nagyobb, mint a $[0, T)$ intervallum.

A 3.3. ábrán az ún. alapsávi nem koherens illesztett szűrős detektor látható.

Teljesen hasonló logikával felépíthető az ún. vivősávi nem koherens illesztett szűrős detektor is, hiszen nyilvánvaló, hogy az

$$\int_0^T r(\tau) \sqrt{2} z_k(\tau) \cos(\omega_0 \tau) d\tau$$

és

$$\int_0^T r(\tau) \sqrt{2} z_k(\tau) \sin(\omega_0 \tau) d\tau$$

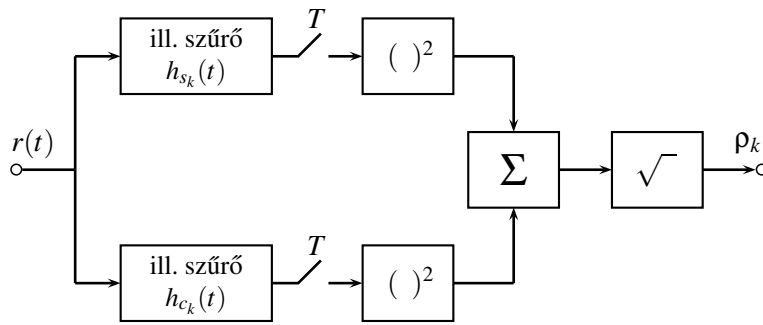
korrelációs műveleteket is helyettesíteni lehet illesztett szűréssel, ha bevezetjük a

$$h_{c_k} = \begin{cases} \sqrt{2} z_k(T-t) \cos(\omega_0(T-t)), & t \in [0, T) \\ 0 & \text{egyébként} \end{cases}$$

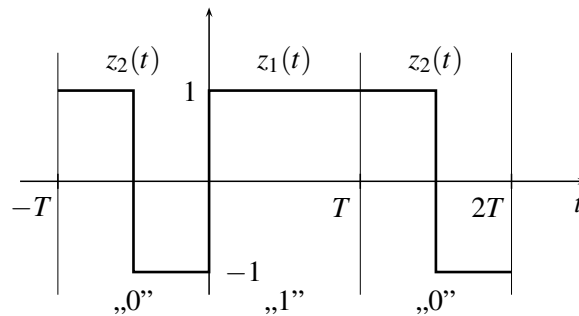
és

$$h_{s_k} = \begin{cases} \sqrt{2} z_k(T-t) \sin(\omega_0(T-t)), & t \in [0, T) \\ 0 & \text{egyébként} \end{cases}$$

súlyfüggvényű szűrőket, és felépítjük velük a 3.4. ábrán látható vevőt.



3.4. ábra. Vivősávi nem koherens illesztett szűrős vevő



3.5. ábra. A küldött jel bináris elemi jelek esetén

3.3. A nem koherens optimális vevők működésének illusztratív összehasonlítása

Zajmentes vizsgálatot és bináris rendszert tételezünk fel az alábbi alapsávi elemi jelekkel:

$$z_1(t) = \begin{cases} 1, & t \in [0, T) \\ 0 & \text{egyébként} \end{cases} \quad z_2(t) = \begin{cases} 1, & t \in [0, \frac{T}{2}) \\ -1, & t \in [\frac{T}{2}, T) \\ 0 & \text{egyébként} \end{cases}$$

Ekkor a küldött jel a 3.5. ábrán látható.

A kvadrátúra korrelációs detektor jelei

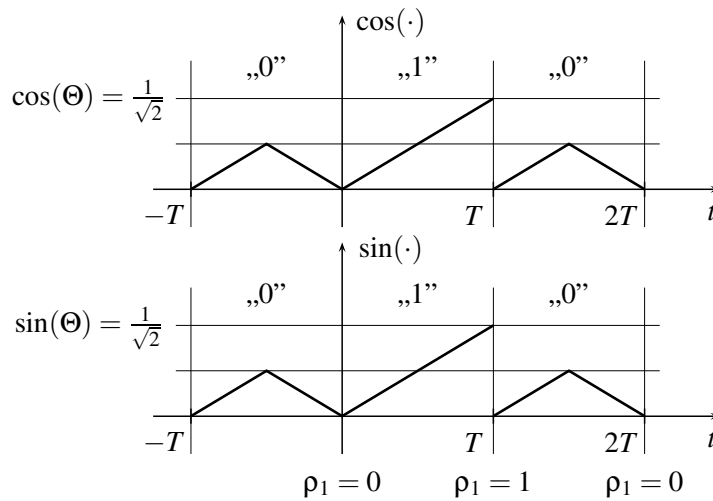
A $z_1(t)$ -hez tartozó korrelátor kimenő jeleit a 3.6. ábrán tüntettük fel $\Theta = 45^\circ$ esetén.

A kvadrátúra korrelációs detektor és az alapsávi nem koherens illesztett szűrős detektor kimenő jeleinek összehasonlítása

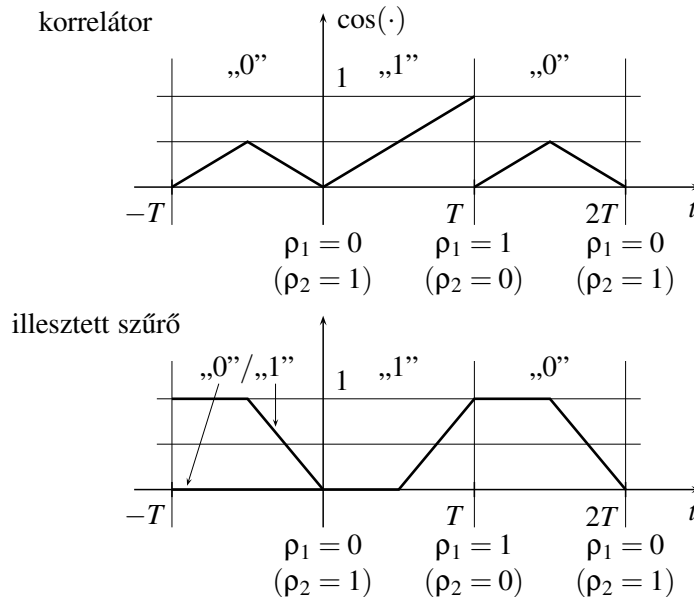
A $z_1(t)$ -hez tartozó detektorok mintavevő előtti kimenő jelei láthatóak a 3.7. ábrán $\Theta = 0^\circ$ esetén. A $\Theta = 0^\circ$ miatt a szinuszos oldal kimenő jele azonosan nulla, így elegendő a koszinuszos oldalt vizsgálni. Megjegyzendő, hogy ekkor a rendszer a koherens optimális detektorhoz hasonlóan viselkedik a $(\cdot)^2$ és $\sqrt{\cdot}$ műveletek nélkül. A mintavételi időpontokban a jelek azonos értékűek.

A vivősávi nem koherens illesztett szűrős detektor kimenő jeleinek illusztrálása

A $z_1(t)$ -hez tartozó detektorok mintavevő előtti kimenő jeleit a 3.8. ábrán tüntettük fel $\Theta = 45^\circ$



3.6. ábra. A kvadratura korrelációs detektor jelei

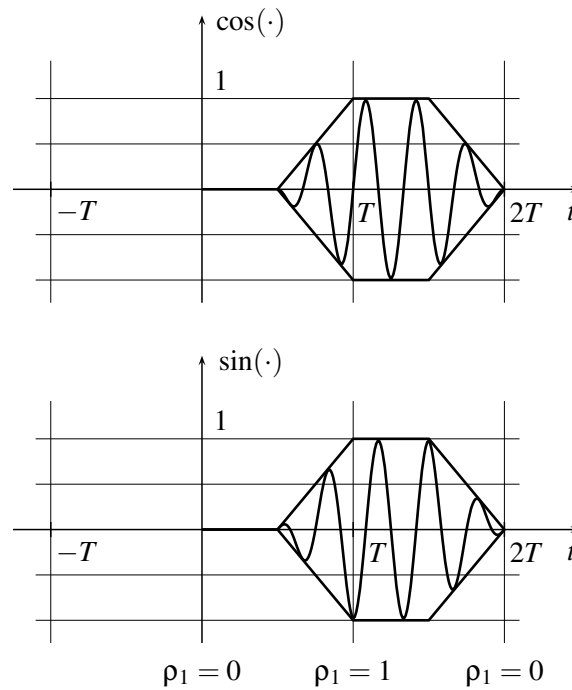
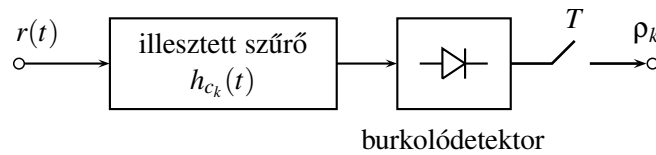


3.7. ábra. A kvadratura korrelációs és az illesztett szűrős detektor összehasonlítása

esetén. Ha a szűrő súlyfüggvénye

$$h_{c_k} = \sqrt{2}z_k(T-t) \cos(\omega_0(T-t))$$

alakú, akkor a szűrő sávszűrő típusú a vivőfrekvencia körüli átviteli sávval. Éppen ezért a szűrő kimenő jele most vivőfrekvenciás komponenseket is tartalmaz, de a két szűrő kimenetén a vivő fázisai között éppen 90° eltérés van. Amiből világosan következik, hogy a vivősávi detektor kialakításához elegendő egy illesztett szűrőt megvalósítani, hiszen a vett jel burkolója mindkét szűrő kimenetén hordozza a szükséges információt. Az egyszerűsített vivősávi detektor felépítése a 3.9. ábrán látható. A megoldás alapvető előnye, hogy nincs szükség a kvadratura komponensek egyidejű kezelésére.

3.8. ábra. A vivősávi nem koherens illesztett szűrős detektor kimenő jelei ($\omega_0 T = 2k\pi$)

3.9. ábra. Az egyszerűsített vivősávi detektor felépítése

3.4. A nem koherens rendszerek hibavalószínűsége

Azonos jelenergiákat és azonos a priori valószínűségeket tételezünk fel (vagyis $E_k = E$ és $\pi_k = \frac{1}{M}$), az elemi jelek ortogonálisak, azaz

$$z_k(t) = \sqrt{E} \varphi_k(t), \quad k = 1, \dots, M, \quad N = M.$$

Az optimális detektorban a döntés a

$$\hat{k} = \underset{k}{\operatorname{argmax}} \{\rho_k\}$$

szabály szerint történik.

A hibavalószínűséget az alábbi módon definiáljuk:

$$\begin{aligned} P_{e_k} &= \mathbf{P}(\hat{k} \neq k | \mathbf{x}_k) = \\ &= \mathbf{P}(\text{nem } k\text{-ra döntünk, ha } \mathbf{x}_k(\mathbf{z}_k) \text{ volt az üzenet}) = \\ &= 1 - \mathbf{P}(k\text{-ra döntünk, ha } \mathbf{x}_k(\mathbf{z}_k) \text{ volt az üzenet}) = \\ &= 1 - P_{c_k}, \end{aligned}$$

ahol P_{c_k} a helyes döntés valószínűsége feltéve, hogy a k -adik üzenetet küldték. A döntési szabály

alkalmazásával

$$\begin{aligned} P_{c_k} &= \mathbf{P}(\rho_k > \rho_m \forall m \neq k \mid \mathbf{x}_k) = \\ &= \mathbf{P}(\rho_k > \rho_1, \dots, \rho_k > \rho_{k-1}, \rho_k > \rho_{k+1}, \dots, \rho_k > \rho_M \mid \mathbf{x}_k). \end{aligned}$$

A hibaarány számításához ismerni kell a $\mathbf{p} = (\rho_1, \dots, \rho_M)$ együttes feltételes valószínűségi sűrűségfüggvényét. Tudjuk, hogy ha a k -adik üzenetet küldték, akkor

$$\mathbf{z}_k = (\overset{(1)}{0}, \overset{(2)}{0}, \dots, \overset{(k-1)}{0}, \overset{(k)}{\sqrt{E}}, \overset{(k+1)}{0}, \dots, \overset{(M)}{0}),$$

$$\mathbf{r} = (\mathbf{r}_c, \mathbf{r}_s) = (\mathbf{v}_{c_1}, \dots, \cos(\Theta)\sqrt{E} + \mathbf{v}_{c_k}, \dots, \mathbf{v}_{c_M}, \mathbf{v}_{s_1}, \dots, \sin(\Theta)\sqrt{E} + \mathbf{v}_{s_k}, \dots, \mathbf{v}_{s_M}),$$

így az l -edik detektor kimenetén az alábbi jelet kapjuk:

$$\begin{aligned} \rho_l &= \sqrt{\langle \mathbf{r}_c, \mathbf{z}_l \rangle^2 + \langle \mathbf{r}_s, \mathbf{z}_l \rangle^2} = \\ &= \sqrt{\langle \sqrt{E} r_{c_l} \rangle^2 + \langle \sqrt{E} r_{s_l} \rangle^2} = \\ &= \sqrt{E} \sqrt{r_{c_l}^2 + r_{s_l}^2} = \\ &= \begin{cases} \sqrt{E} \sqrt{(\cos(\Theta)\sqrt{E} + \mathbf{v}_{c_k})^2 + (\sin(\Theta)\sqrt{E} + \mathbf{v}_{s_k})^2}, & l = k \\ \sqrt{E} \sqrt{\mathbf{v}_{c_l}^2 + \mathbf{v}_{s_l}^2}, & l \neq k \end{cases} \end{aligned}$$

amelyből világosan látszik, hogy zajmentes esetben csak a k -adik detektor kimenetén jelenik meg nullától különböző jel (akkor, ha a k -adik üzenetet küldték).

A korábbiakban meghatároztuk az \mathbf{r} feltételes valószínűségi sűrűségfüggvényét, miszerint

$$p_{\mathbf{r}}(\mathbf{y} \mid \mathbf{x}_k) = p_{\mathbf{r}}(\mathbf{y}_c, \mathbf{y}_s \mid \mathbf{x}_k) = \left(\frac{1}{\pi N_0} \right)^N \exp\left(-\frac{\|\mathbf{y}\|^2}{N_0}\right) \exp\left(-\frac{E}{N_0}\right) I_0\left(2\frac{\rho_k}{N_0}\right),$$

amit esetünkben az alábbi alakra hozhatunk:

$$p_{\mathbf{r}}(\mathbf{y} \mid \mathbf{x}_k) = \exp\left(-\frac{E}{N_0}\right) I_0\left(2\frac{\sqrt{y_{c_k}^2 + y_{s_k}^2}}{N_0}\sqrt{E}\right) \prod_{j=1}^N \frac{1}{\pi N_0} \exp\left(-\frac{y_{c_j}^2 + y_{s_j}^2}{N_0}\right).$$

Vezessünk be egy valószínűségiváltozó-rendszert

$$\begin{aligned} \mathbf{R} &= \frac{\mathbf{p}}{\sqrt{E}} = (R_1, \dots, R_N), & R_l &= \sqrt{r_{c_l}^2 + r_{s_l}^2} \\ \mathbf{\Phi} &= (\Phi_1, \dots, \Phi_N), & \Phi_l &= \arctg \frac{r_{s_l}}{r_{c_l}}, \end{aligned}$$

amiből

$$p_{\mathbf{R}, \mathbf{\Phi}}(\mathbf{r}, \mathbf{\Phi} \mid \mathbf{x}_k) = \frac{1}{\pi N_0} r_k \exp\left(-\frac{E}{N_0}\right) \exp\left(-\frac{r_k^2}{N_0}\right) I_0\left(2r_k \frac{\sqrt{E}}{N_0}\right) \prod_{\substack{l=1 \\ l \neq k}}^N \frac{1}{\pi N_0} r_l \exp\left(-\frac{r_l^2}{N_0}\right).$$

Az új valószínűségi vektorváltozók sűrűségfüggvényének alakja azért ilyen, mert általában igaz a következő. Ha adott egy X, Y valószínűségi változó páros, és előállítjuk ezek egy függvénytranszformációját, például az

$$\begin{aligned} R &= f(X, Y) = \sqrt{X^2 + Y^2}, & r &= f(x, y) = \sqrt{x^2 + y^2} \\ \Phi &= g(X, Y) = \arctg \frac{Y}{X}, & \varphi &= g(x, y) = \arctg \frac{y}{x} \end{aligned}$$

függvényekkel, akkor az X, Y együttes sűrűségfüggvényének ismeretében az R, Φ együttes sűrűségfüggvényét az alábbi eljárással határozhatjuk meg:

$$p_{R,\Phi}(r, \varphi) = p_{X,Y}(x, y) \left| \frac{\partial(x, y)}{\partial(r, \varphi)} \right|, \quad \begin{array}{l} x = r \cos(\varphi) \\ y = r \sin(\varphi) \end{array}$$

vagy

$$p_{X,Y}(x, y) = p_{R,\Phi}(r, \varphi) \left| \frac{\partial(r, \varphi)}{\partial(x, y)} \right|,$$

ahol

$$\left| \frac{\partial(x, y)}{\partial(r, \varphi)} \right| = J = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} \\ \frac{\partial x}{\partial \varphi} & \frac{\partial y}{\partial \varphi} \end{vmatrix} = \begin{vmatrix} \cos(\varphi) & \sin(\varphi) \\ -r \sin(\varphi) & r \cos(\varphi) \end{vmatrix} = r$$

a Jacobi-mátrix determinánsa.

Esetünkben tehát

$$p_{R,\Phi}(r, \varphi) = p_{X,Y}(x, y) \cdot r = p_{X,Y}(r \cos(\varphi), r \sin(\varphi)) \cdot r,$$

így ha például

$$p_{X,Y}(x, y) = \frac{1}{\pi N_0} \exp\left(-\frac{x^2 + y^2}{N_0}\right),$$

akkor

$$p_{R,\Phi}(r, \varphi) = \frac{1}{\pi N_0} r \exp\left(-\frac{r^2}{N_0}\right),$$

illetve, ha

$$p_{X,Y}(x, y) = \frac{1}{\pi N_0} \exp\left(-\frac{E}{N_0}\right) \exp\left(-\frac{x^2 + y^2}{N_0}\right) I_0\left(2 \frac{\sqrt{x^2 + y^2}}{N_0} \sqrt{E}\right),$$

akkor

$$p_{R,\Phi}(r, \varphi) = \frac{1}{\pi N_0} r \exp\left(-\frac{E}{N_0}\right) \exp\left(-\frac{r^2}{N_0}\right) I_0\left(2 \frac{r}{N_0} \sqrt{E}\right).$$

Tudjuk továbbá, hogy

$$p_X(x) = \int_{-\infty}^{\infty} p_{X,Y}(x, y) dy.$$

A hibavalószínűség számításához szükséges $p_{\mathbf{R}}(\mathbf{r} | \mathbf{x}_k)$ sűrűségfüggvényt ennek alapján a

$$p_{\mathbf{R}}(\mathbf{r} | \mathbf{x}_k) = \frac{2}{N_0} r_k \exp\left(-\frac{E}{N_0}\right) \exp\left(-\frac{r_k^2}{N_0}\right) I_0\left(2 r_k \frac{\sqrt{E}}{N_0}\right) \prod_{\substack{l=1 \\ l \neq k}}^N \frac{2}{N_0} r_l \exp\left(-\frac{r_l^2}{N_0}\right)$$

alakban adhatjuk meg. Jól látszik, hogy az \mathbf{R} (illetve a \mathbf{p}) vektor egyes komponensei független valószínűségi változók, hiszen sűrűségfüggvényeik szorzat alakban írhatók fel.

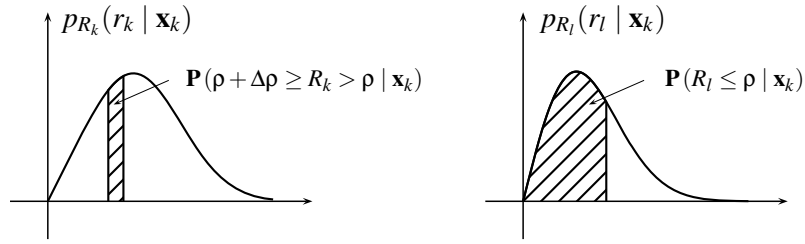
A k -adik üzenet küldése esetén a k -adik nem koherens detektor kimenetén lévő $R_k = \frac{p_k}{\sqrt{E}}$ jel feltételes valószínűségi sűrűségfüggvénye

$$p_{R_k}(r_k | \mathbf{x}_k) = \frac{2}{N_0} r_k \exp\left(-\frac{E}{N_0}\right) \exp\left(-\frac{r_k^2}{N_0}\right) I_0\left(2 r_k \frac{\sqrt{E}}{N_0}\right)$$

alakú. Ezt Rice-eloszlásnak nevezzük.

A k -adik üzenet küldése esetén az l -edik ($l \neq k$) nem koherens detektor kimenetén lévő $R_l = \frac{p_l}{\sqrt{E}}$ jel feltételes valószínűségi sűrűségfüggvénye pedig

$$p_{R_l}(r_l | \mathbf{x}_k) = \frac{2}{N_0} r_l \exp\left(-\frac{r_l^2}{N_0}\right)$$



3.10. ábra. Optimális döntés a Rice- és Rayleigh-eloszlások alapján

alakú. Ezt Rayleigh-eloszlásnak nevezzük.

Egészében elmondhatjuk, hogy a k -adik üzenet küldése esetén (optimális döntést feltételezve) a következő történik. A k -adik detektor kimenetén lévő Rice-eloszlású R_k független valószínűségi változó és a többi $(M-1)$ detektor kimenetén lévő Rayleigh-eloszlású $\{R_l\}, l \neq k$ valószínűségi változók megfigyelése után a maximális értékű indexre döntünk.

Ez alapján a feladat az, hogy határozzuk meg a P_{c_k} értékét, ami az alábbiak szerint tehető meg.

$$P_{c_k} = \mathbf{P}(R_k > R_l \quad \forall l \neq k \mid \mathbf{x}_k)$$

Tudjuk, hogy adott ρ esetén ($l \neq k$)-ra

$$\begin{aligned} \mathbf{P}(\rho \geq R_l \mid \mathbf{x}_k) &= 2 \int_0^\rho \frac{r_l}{N_0} \exp\left(-\frac{r_l^2}{N_0}\right) dr_l = \\ &= \left[-\exp\left(-\frac{r_l^2}{N_0}\right) \right]_0^\rho = \\ &= 1 - \exp\left(-\frac{\rho^2}{N_0}\right) \end{aligned}$$

$$\mathbf{P}(\rho + \Delta\rho \geq R_k > \rho \mid \mathbf{x}_k) = 2 \frac{\rho}{N_0} \exp\left(-\frac{E+\rho^2}{N_0}\right) I_0\left(2\rho \frac{\sqrt{E}}{N_0}\right) \Delta\rho$$

Ily módon a függetlenséget felhasználva

$$\begin{aligned} \mathbf{P}(R_k > R_1, \dots, R_k > R_{k-1}, \rho + \Delta\rho \geq R_k > \rho, R_k > R_{k+1}, \dots, R_k > R_M \mid \mathbf{x}_k) = \\ = 2 \frac{\rho}{N_0} \exp\left(-\frac{E+\rho^2}{N_0}\right) I_0\left(2\rho \frac{\sqrt{E}}{N_0}\right) \left(1 - \exp\left(-\frac{\rho^2}{N_0}\right)\right)^{M-1} \Delta\rho, \end{aligned}$$

tehát

$$P_{c_k} = \int_0^\infty 2 \frac{\rho}{N_0} \exp\left(-\frac{E+\rho^2}{N_0}\right) I_0\left(2\rho \frac{\sqrt{E}}{N_0}\right) \left(1 - \exp\left(-\frac{\rho^2}{N_0}\right)\right)^{M-1} d\rho,$$

és $P_{e_k} = 1 - P_{c_k}$, $P_{e_k} = P_e$, $P_{c_k} = P_c$. A feladatot a 3.10. ábra illusztrálja ($M = 2$).

3.5. Példa a nem koherens rendszerek hibaarányának számítására

Bináris eset

$$M = 2, N = 2$$

$$\begin{aligned} x_1(t) &= \sqrt{2E} \sqrt{\frac{2}{T}} \cos\left(\frac{\pi t}{T}\right) \cos(\omega_0 t), & \phi_1 &= \sqrt{\frac{2}{T}} \cos\left(\frac{\pi t}{T}\right), \\ x_2(t) &= \sqrt{2E} \sqrt{\frac{2}{T}} \sin\left(\frac{\pi t}{T}\right) \cos(\omega_0 t), & \phi_2 &= \sqrt{\frac{2}{T}} \sin\left(\frac{\pi t}{T}\right), \end{aligned}$$

ahol $t \in [0, T)$. Az új ortonormált bázis:

$$\begin{aligned}\Psi_{c_1}(t) &= \sqrt{2}\sqrt{\frac{2}{T}} \cos\left(\frac{\pi t}{T}\right) \cos(\omega_0 t), \\ \Psi_{c_2}(t) &= \sqrt{2}\sqrt{\frac{2}{T}} \sin\left(\frac{\pi t}{T}\right) \cos(\omega_0 t), \\ \Psi_{s_1}(t) &= \sqrt{2}\sqrt{\frac{2}{T}} \cos\left(\frac{\pi t}{T}\right) \sin(\omega_0 t), \\ \Psi_{s_2}(t) &= \sqrt{2}\sqrt{\frac{2}{T}} \sin\left(\frac{\pi t}{T}\right) \sin(\omega_0 t).\end{aligned}$$

Ezért

$$\begin{aligned}x_1(t) &= \sqrt{E}\Psi_{c_1}(t), & \mathbf{x}_1 &= (\sqrt{E}, 0), \\ x_2(t) &= \sqrt{E}\Psi_{c_2}(t), & \mathbf{x}_2 &= (0, \sqrt{E}).\end{aligned}$$

$$\begin{aligned}P_c &= \int_0^\infty 2\frac{\rho}{N_0} \exp\left(-\frac{E+\rho^2}{N_0}\right) I_0\left(2\rho\frac{\sqrt{E}}{N_0}\right) \left(1 - \exp\left(-\frac{\rho^2}{N_0}\right)\right) d\rho = \\ &= 1 - \int_0^\infty 2\frac{\rho}{N_0} \exp\left(-\frac{E+2\rho^2}{N_0}\right) I_0\left(2\rho\frac{\sqrt{E}}{N_0}\right) d\rho,\end{aligned}$$

mivel az

$$\int_0^\infty 2\frac{\rho}{N_0} \exp\left(-\frac{E+\rho^2}{N_0}\right) I_0\left(2\rho\frac{\sqrt{E}}{N_0}\right) d\rho = 1,$$

ugyanis az integrálandó függvény valószínűségi sűrűségfüggvény.

Az összefüggés alapján a hibaarányt a

$$P_e = \int_0^\infty 2\frac{\rho}{N_0} \exp\left(-\frac{E+2\rho^2}{N_0}\right) I_0\left(2\rho\frac{\sqrt{E}}{N_0}\right) d\rho$$

integrál kiértékelésével kaphatjuk meg. Ez zárt alakban előállítható:

$$P_e = \frac{1}{2} \exp\left(-\frac{E}{2N_0}\right).$$

Az integrálok analitikus kiértékelése

A hibaarány számításához szükség van a

$$P_c = \int_0^\infty 2\frac{\rho}{N_0} \exp\left(-\frac{E+\rho^2}{N_0}\right) I_0\left(2\rho\frac{\sqrt{E}}{N_0}\right) \left(1 - \exp\left(-\frac{\rho^2}{N_0}\right)\right)^{M-1} d\rho$$

integrál kiértékelésére. Bevezetve az $x = \frac{\rho^2}{N_0}$, $\frac{dx}{d\rho} = 2\frac{\rho}{N_0}$ új integrálási változót, a

$$\begin{aligned}P_c &= \int_0^\infty \exp\left(-\frac{E}{N_0} - x\right) I_0\left(2\sqrt{x}\frac{\sqrt{E}}{N_0}\right) (1 - \exp(-x))^{M-1} dx = \\ &= \int_0^\infty \exp\left(-\frac{E}{N_0} - x\right) I_0\left(2\sqrt{x}\frac{\sqrt{E}}{N_0}\right) \sum_{m=0}^{M-1} (-1)^m \binom{M-1}{m} \exp(-mx) dx = \\ &= \sum_{m=0}^{M-1} (-1)^m \binom{M-1}{m} \exp\left(-\frac{E}{N_0}\right) \int_0^\infty \exp(-(m+1)x) I_0\left(2\sqrt{x}\frac{\sqrt{E}}{N_0}\right) dx\end{aligned}$$

kifejezéshez jutunk, azaz a kulcskérdés az

$$\int_0^{\infty} \exp(-(m+1)x) I_0 \left(2\sqrt{x \frac{E}{N_0}} \right) dx$$

integrál kiértékelése.

Matematikai kézikönyvekből ismert, hogy léteznek az alábbi Laplace-transzformált:

$$\mathcal{L} \left\{ \left(\frac{t}{k} \right)^{\frac{\mu-1}{2}} I_{\mu-1} (2\sqrt{kt}) \right\} = \frac{1}{s^{\mu}} \exp \left(\frac{k}{s} \right), \quad \mu = 1, 2, \dots$$

és $\mu = 1$ esetén

$$\mathcal{L} \left\{ I_0 (2\sqrt{kt}) \right\} = \frac{1}{s} \exp \left(\frac{k}{s} \right).$$

Tudjuk továbbá, hogy

$$\begin{aligned} \mathcal{L} \left\{ \exp(-ct) I_0 (2\sqrt{kt}) \right\} &= \frac{1}{s+c} \exp \left(\frac{k}{s+c} \right) = \\ &= \int_0^{\infty} \exp(-ct) I_0 (2\sqrt{kt}) \exp(-st) dt. \end{aligned}$$

Ez utóbbi egyenlet felhasználásával a keresett integrál kiszámítható, ha a kifejezésbe a $k \equiv \frac{E}{N_0}$, $c \equiv m+1$, $s \equiv 0$, $t \equiv x$ értékeket helyettesítjük, mivel ekkor

$$\int_0^{\infty} \exp(-(m+1)x) I_0 \left(2\sqrt{x \frac{E}{N_0}} \right) dx = \frac{1}{m+1} \exp \left(\frac{1}{m+1} \frac{E}{N_0} \right).$$

Ezt az eredményt behelyettesítve a korábbi összefüggésekbe, a P_c értéke a

$$\begin{aligned} P_c &= \sum_{m=0}^{M-1} (-1)^m \binom{M-1}{m} \exp \left(-\frac{E}{N_0} \right) \frac{1}{m+1} \exp \left(\frac{1}{m+1} \frac{E}{N_0} \right) = \\ &= \sum_{m=0}^{M-1} (-1)^m \binom{M-1}{m} \frac{1}{m+1} \exp \left(-\frac{E}{N_0} \frac{m}{m+1} \right) = \\ &= 1 + \sum_{m=1}^{M-1} (-1)^m \binom{M-1}{m} \frac{1}{m+1} \exp \left(-\frac{E}{N_0} \frac{m}{m+1} \right) = \\ &= 1 - P_e \end{aligned}$$

kifejezéssel számolható, amelyből a hibaarányra

$$P_e = \sum_{m=1}^{M-1} (-1)^{m+1} \binom{M-1}{m} \frac{1}{m+1} \exp \left(-\frac{E}{N_0} \frac{m}{m+1} \right)$$

érték adódik.

Meg kell jegyezni, hogy nagy M -ek esetére és kis hibaarányok mellett a váltakozó előjelek miatt a kifejezés toleranciaérzékeny, de például bináris esetben egyszerűen kiértékelhető, ugyanis $M = 2$ esetén

$$P_e = \frac{1}{2} \exp \left(-\frac{E}{2N_0} \right).$$

4. fejezet

A koherens és nem koherens átviteli rendszerek összehasonlítása

Az összehasonlítást (hibaarány alapján) az alábbi közös feltételekkel végezzük el:

- $E_k = E$, $\pi_k = \frac{1}{M}$
- Az elemi jelek ortogonálisak, azaz $N = M$.
- Optimális vevőstruktúrák.

4.1. Koherens csatorna

$$x_k(t) = \sqrt{E} \varphi_k(t), \quad k = 1, 2, \dots, M$$

$$\mathbf{x}_k = (\overset{(1)}{0}, \overset{(2)}{0}, \dots, \overset{(k-1)}{0}, \overset{(k)}{\sqrt{E}}, \overset{(k+1)}{0}, \dots, \overset{(M)}{0})$$

$\mathbf{r} = \mathbf{x}_k + \mathbf{v}$, ha a k -adik üzenetet küldték.

$$p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k) = \left(\frac{1}{\sqrt{\pi N_0}} \right)^M \exp \left(-\frac{(y_k - \sqrt{E})^2}{N_0} \right) \prod_{\substack{m=1 \\ m \neq k}}^M \exp \left(-\frac{y_m^2}{N_0} \right)$$

Tudjuk, hogy az optimális vevőben előállítjuk a

$$\rho_l = \langle \mathbf{r}, \mathbf{x}_l \rangle = \sqrt{E} r_l = \begin{cases} \sqrt{E}(\sqrt{E} + v_k), & l = k \\ \sqrt{E} v_l, & l \neq k \end{cases}$$

skalárszorzatokat, és a döntés a

$$\hat{k} = \underset{k}{\operatorname{argmax}} \{ \rho_k \} = \underset{k}{\operatorname{argmax}} \{ R_k \} = \underset{k}{\operatorname{argmax}} \{ r_k \}$$

alapján történik, ahol $R_k = \frac{\rho_k}{\sqrt{E}} = r_k$.

A $p_{\mathbf{r}}(\mathbf{y} | \mathbf{x}_k)$ szerkezetéből jól látható, hogy az \mathbf{r} vektor egyes komponensei független Gauss-eloszlású valószínűségi változók. A P_c pedig most is a

$$\begin{aligned} P_c = P_{c_k} &= \mathbf{P}(R_k > R_m, \forall m \neq k | \mathbf{x}_k) = \\ &= \int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{\pi N_0}} \right)^M \exp \left(-\frac{(\rho - \sqrt{E})^2}{N_0} \right) \left(\int_{-\infty}^{\rho} \exp \left(-\frac{y^2}{N_0} \right) dy \right)^{M-1} d\rho \end{aligned}$$

egyenletből határozható meg, amely átalakítások után a

$$\begin{aligned}
P_c &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{(\rho - \sqrt{E})^2}{N_0}\right) \left(\int_{-\infty}^{\rho} \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{y^2}{N_0}\right) dy \right)^{M-1} d\rho = \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{(\rho - \sqrt{E})^2}{N_0}\right) \left(1 - \int_{\rho}^{\infty} \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{y^2}{N_0}\right) dy \right)^{M-1} d\rho = \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{(\rho - \sqrt{E})^2}{N_0}\right) \left(1 - Q\left(\frac{\rho}{\sqrt{N_0}} \sqrt{2}\right) \right)^{M-1} d\rho = \\
&= 1 - P_e
\end{aligned}$$

formában adható meg. Bináris esetben

$$P_e = \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{(\rho - \sqrt{E})^2}{N_0}\right) Q\left(\frac{\rho}{\sqrt{N_0}} \sqrt{2}\right) d\rho = Q\left(\sqrt{\frac{E}{N_0}}\right).$$

4.2. Nem koherens csatorna

$$\begin{aligned}
x_k(t) &= \sqrt{E} \Psi_{c_k}(t), \quad k = 1, 2, \dots, M \\
z_k(t) &= \sqrt{E} \Phi_k(t), \quad k = 1, 2, \dots, M
\end{aligned}$$

$$\mathbf{r} = (\mathbf{r}_c, \mathbf{r}_s) = (\cos(\Theta) \mathbf{x}_k + \mathbf{v}_c, \sin(\Theta) \mathbf{x}_k + \mathbf{v}_s)$$

$$p_{\mathbf{r}}(\mathbf{y}_c, \mathbf{y}_s | \mathbf{x}_k) = \left(\frac{1}{\pi N_0}\right)^N \exp\left(-\frac{\|\mathbf{y}\|^2}{N_0}\right) \exp\left(-\frac{E_k}{N_0}\right) I_0\left(2\frac{\rho_k}{N_0}\right)$$

Tudjuk, hogy a korrekt vétel valószínűségét a

$$P_c = \int_0^{\infty} 2 \frac{\rho}{N_0} \exp\left(-\frac{E + \rho^2}{N_0}\right) I_0\left(2\rho \frac{\sqrt{E}}{N_0}\right) \left(1 - \exp\left(-\frac{\rho^2}{N_0}\right)\right)^{M-1} d\rho$$

kifejezés adja meg.

A 4.1. ábrán a hibavalószínűségeket adtuk meg az $\frac{E_k}{N_0}$ függvényében. Látható, hogy $M = N \rightarrow \infty$ esetén mind a koherens, mind a nem koherens rendszer hibavalószínűsége a

$$\lim_{M \rightarrow \infty} P_e = \begin{cases} 1, & \text{ha } \frac{E}{N_0} \frac{1}{\log_2 M} \leq \ln 2 \\ 0, & \text{ha } \frac{E}{N_0} \frac{1}{\log_2 M} > \ln 2 \end{cases}$$

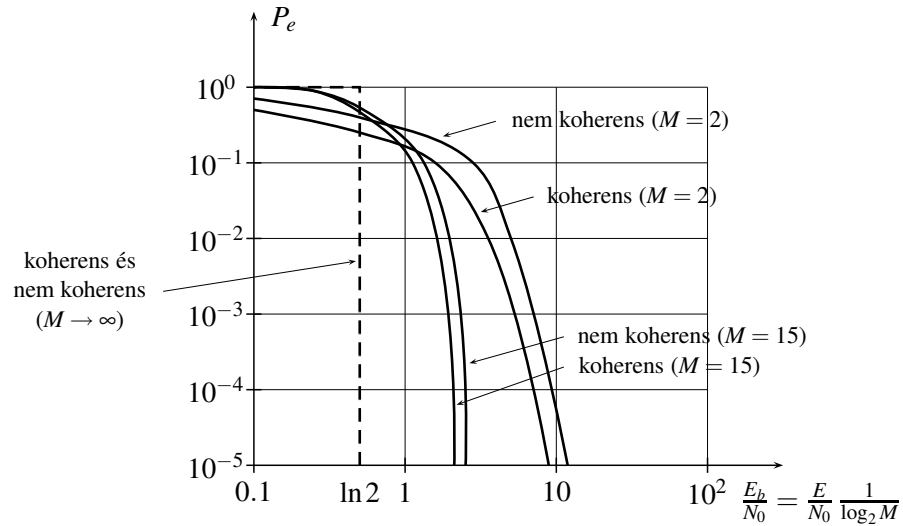
határfüggvényhez tart. A határon

$$\frac{E}{N_0} \frac{1}{\log_2 M} = \ln 2,$$

és bevezetve az $R = \frac{\log_2 M}{T}$ átviteli sebességet, valamint a $P = \frac{E}{T}$ teljesítményt

$$\frac{E}{N_0} \frac{1}{\log_2 M} \frac{T}{T} = \frac{1}{R} \frac{P}{N_0} = \ln 2,$$

$$R = \frac{P}{N_0} \log_2 e.$$

4.1. ábra. Hibavalószínűségek az $\frac{E_b}{N_0}$ függvényében

Így a feltétel a következő

$$\lim_{M \rightarrow \infty} P_e = \begin{cases} 1, & \text{ha } R \geq \frac{P}{N_0} \log_2 e \\ 0, & \text{ha } R < \frac{P}{N_0} \log_2 e \end{cases}$$

amiből nyilvánvaló, hogy

$$C_\infty = \frac{P}{N_0} \log_2 e$$

a csatorna kapacitása, ha a sávszélesség minden határon túl nő, hiszen Shannon eredményei alapján

$$\begin{aligned} C_\infty &= \lim_{W \rightarrow \infty} W \log_2 \left(1 + \frac{P}{N_0 W} \right) = \\ &= \lim_{W \rightarrow \infty} \log_2 e \ln \left(1 + \frac{P}{N_0 W} \right)^W = \\ &= \log_2 e \cdot \frac{P}{N_0}. \end{aligned}$$

5. fejezet

A vektortér kiterjesztése a sávhatárolt jelekre

Legyen a forrás által előállított szimbólumok sorozata $\{\xi_{k_i}\}$, amelyre igaz, hogy minden eleme egy M értékészletű ábécéből veszi fel az értékeit, azaz $k_i \in \{1, \dots, k, \dots, M\}$, és az i index az időbeli sorrendet jelöli. (Ha az eredeti forrás L értékészletű szimbólumsorozatot állít elő, és azokból K darabot összefogva hozzuk létre a ξ szimbólumokat, akkor $M = L^K$). Jelöljük a szimbólumidőt T -vel, így módon az i -edik időrésben a ξ_{k_i} szimbólumot visszük át, azaz, ha $k_i = h$, akkor a h -adik szimbólumot.

A $\{\xi_{k_i}\}$ szimbólumsorozat átviteléhez időben folytonos sávkorlátozott függvényeket használunk, ezeket $\{x_{k_i}(t)\}$ -vel jelöljük. A moduláció tehát egyszerű leképezés, miszerint a ξ_{k_i} szimbólumhoz az $x_{k_i}(t - iT)$ időben eltolts elemi jelet rendeljük hozzá.

Tételezzük fel, hogy az elemi jelek sáv szélessége $B = [0, W)$, azaz a jel egyoldalas spektruma csak ebben a frekvenciatartományban különbözik nullától, így az elemi jelek $X_k(f)$ Fourier-transzformáltja

$$X_k(f) = \begin{cases} \text{tetszőleges,} & f \in (-W, W) \\ 0, & f \notin (-W, W) \end{cases},$$

vagyis az elemi jelek

$$X_k(f) = \int_{-\infty}^{+\infty} x_k(t) \exp(-j2\pi ft) dt$$

Fourier-transzformáltjának a tartója a $(-W, W)$ tartomány. (Természetesen ez a megkötés sem teljesül általában, de ez a tárgyalás a valóságos rendszerek működését jól leírja, hiszen a valóságos modulált jelek sáv szélessége értelemszerűen korlátozott frekvenciasávot foglal el).

A vizsgált additív fehér Gauss-zajos csatornában a vevő bemenetére az

$$r_k(t) = x_k(t) + v(t)$$

eredő jel érkezik, ha az adóban éppen a k -dik üzenetet küldték, ahol $x_k(t)$ a hasznos jel, $v(t)$ pedig a fehér Gauss-zaj időfüggvénye.

Tételezzük fel, hogy az $\{x_{k_i}(t)\}$ elemi jelek teljesítménye korlátozott, és P az átlagos teljesítmény. Ilyenkor fehér Gauss-zajos csatornában, ahol $S_n(f) = N_0/2$ a zaj kétoldalas teljesítménysűrűsége, a csatornában mérhető átlagos jel-zaj viszony az

$$SNR = \frac{P}{N_0 W}$$

formában adható meg.

5.1. A négyzetesen integrálható függvények L_2 tere

A négyzetesen integrálható $\{x_k(t)\}$ függvények úgynevezett L_2 tere, két műveletre, az összeadásra, a valós skalár konstanssal való szorzásra zárt, és benne értelmezhető a

$$\langle x(t), y(t) \rangle = \int x(t)y(t)dt, \text{ ha } x(t), y(t) \in L_2$$

skaláris szorzat.

- Ennek alapján definiálni tudjuk az $\{x_k(t)\}$ függvény normanégyzetét, ami az

$$\|x(t)\|^2 = \langle x(t), x(t) \rangle = \int x^2(t)dt < \infty, \text{ ha } x(t) \in L_2$$

kifejezéssel adható meg.

- Az L_2 térben két függvény euklideszi távolságának a négyzetét a

$$d^2(x(t), y(t)) = \|x(t) - y(t)\|^2 = \int (x(t) - y(t))^2 dt, \text{ ha } x(t), y(t) \in L_2$$

kifejezés adja meg.

- Az L_2 térben két függvény azonos, ha a távolságuk 0, ugyanis ilyenkor a két függvény csak nullmértékű halmazon különbözik egymástól.
- Tudjuk, hogy az L_2 térbeli függvények esetén $\|x(t)\|^2 \geq 0$, és ha $\|x(t)\|^2 = 0$, akkor $x(t) = 0$.
- Az L_2 térhez tartozó függvényeknek létezik az

$$X_k(f) = \int_{-\infty}^{+\infty} x_k(t) \exp(-j2\pi ft) dt$$

Fourier-transzformáltja, és az

$$x_k(t) = \int_{-\infty}^{+\infty} X_k(f) \exp(j2\pi ft) df$$

inverz Fourier-transzformáltja.

- Általában két függvény az L_2 térben azonos, ha legfeljebb véges számú, vagy megszámlálhatóan végtelen számú pontban (nullmértékű halmazon) tér el egymástól. Ez az állítás a jelek Fourier-transzformáltjára is igaz.
- Az L_2 térbeli függvények esetében a függvények és azok Fourier-transzformáltjai között fennáll a Parseval-egyenlőség, amely szerint

$$\langle x(t), y(t) \rangle = \int_{-\infty}^{+\infty} x(t)y(t)dt = \langle X(f), Y(f) \rangle = \int_{-\infty}^{+\infty} X(f)Y^*(f)df, \text{ ha } x(t), y(t) \in L_2,$$

ahol $Y^*(f)$ az $y(t)$ függvény Fourier-transzformáltjának komplex konjugáltja. Ennek alapján igaz, hogy

$$\|x(t)\|^2 = \|X(f)\|^2.$$

5.2. A jeltér általános definíciója

A jeltér az L_2 tér egy $S \subseteq L_2$ altere. A négyzetesen integrálható függvények jellegzetes példái a $[0, T)$ tartójú, vagy az általánosan véges tartójú időfüggvények, de lehetnek ezek a sávkorlátozott, végtelen időbeli tartóval rendelkező jelek is.

A jeltérben a jelek leírásához mindig szükség van egy **ortogonális** $\{\varphi_j(t), j \in I\}$ **bázisra**, ahol I az egész számok indextere, és fennáll a

$$\langle \varphi_j(t), \varphi_k(t) \rangle = \int_{-\infty}^{+\infty} \varphi_j(t) \varphi_k(t) dt = \begin{cases} > 0, & j = k \\ 0, & j \neq k \end{cases}$$

összefüggés, és amely a modulációs (kommunikációs) rendszerben használt $\{x_k(t)\}$ elemi jelekre nézve **teljes**, azaz minden $\{x_k(t)\}$ függvény megadható az

$$x_k(t) = \sum_{j \in I} \frac{\langle x_k(t), \varphi_j(t) \rangle}{\|\varphi_j(t)\|^2} \varphi_j(t)$$

kifejezés szerint a bázisfüggvények lineáris kombinációjaként.

Ha a bázis normált, akkor $\|\varphi_j(t)\|^2 = 1$, és a fenti kifejezés az

$$x_k(t) = \sum_{j \in I} \langle x_k(t), \varphi_j(t) \rangle \varphi_j(t) = \sum_{j \in I} x_{kj} \varphi_j(t)$$

alakban adható meg, ami az $\{x_k(t)\}$ függvények általánosított Fourier-sora a $\{\varphi_j(t), j \in I\}$ ortonormált bázisfüggvények szerint.

Ilyenkor igaz, hogy

$$\langle x_k(t), x_l(t) \rangle = \int_{-\infty}^{+\infty} x_k(t) x_l(t) dt = \langle \mathbf{x}_k, \mathbf{x}_l \rangle = \sum_{j \in I} x_{kj} x_{lj},$$

ahol $\langle \mathbf{x}_k, \mathbf{x}_l \rangle$ az $\mathbf{x}_k = \{x_{k1}, \dots, x_{kj}, \dots\}$ és az $\mathbf{x}_l = \{x_{l1}, \dots, x_{lj}, \dots\}$ vektorok skalárszorzata, és $x_{kj} = \langle x_k(t), \varphi_j(t) \rangle$, $x_{lj} = \langle x_l(t), \varphi_j(t) \rangle$, amiből

$$\|x_k(t)\|^2 = \langle x_k(t), x_k(t) \rangle = \int_{-\infty}^{+\infty} x_k^2(t) dt = \langle \mathbf{x}_k, \mathbf{x}_k \rangle = \sum_{j \in I} x_{kj}^2 = \|\mathbf{x}_k\|^2,$$

ahol $\|\mathbf{x}_k\|$ az \mathbf{x}_k vektor hosszúsága (normája).

5.3. Az alapsávi sávkorlátozott jelek leírása a jeltérben

Legyen az $x(t)$ jel sávszélessége véges ($B = [0, W)$), azaz legyen a jel Fourier-transzformáltjának a tartója az $(-W, W)$ intervallum. Tudjuk, hogy a véges (például $(-W, W)$) tartójú jelek Fourier-sorba fejthetők az

$$\{\exp(-j2\pi f k T), k \in \mathbb{Z}, f \in (-W, W)\}$$

frekvenciában periódikus jelek segítségével, ahol \mathbb{Z} az egész számok halmaza, és $T = 1/2W$. Ezek a jelek ugyanis ortogonális bázist alkotnak, ami egyszerűen belátható az alábbi integrál kiértékelésével:

$$\int_{-W}^{+W} \exp(-j2\pi f k T) \exp(j2\pi f l T) df = \left[\frac{\exp(-j2\pi f(k-l)T)}{-j2\pi(k-l)T} \right]_{-W}^W =$$

$$\begin{aligned}
&= \frac{\exp(j2\pi W(k-l)T) - \exp(-j2\pi W(k-l)T)}{j2\pi(k-l)T} = \frac{1}{T} \frac{\sin(\pi(k-l))}{\pi(k-l)} = \\
&= \begin{cases} \frac{1}{T}, & k=l \\ 0, & k \neq l \end{cases}.
\end{aligned}$$

Ebből látható, hogy a

$$\{\sqrt{T}\exp(-j2\pi fkT), \quad k \in \mathbb{Z}, \quad f \in (-W, W)\}$$

jelkészlet ortonormált bázist alkot minden L_2 térbeli sávhatárolt függvényre. Ennek alapján a jelek Fourier- transzformáltja az

$$X(f) = \sum_{k \in \mathbb{Z}} \langle X(f), \sqrt{T}\exp(-j2\pi fkT) \rangle \sqrt{T}\exp(-j2\pi fkT)$$

általános egyenlettel leírható. Ha ennek az egyenletnek mindkét oldalán végrehajtjuk az inverz Fourier-transzformáció műveletét, akkor az alábbi összefüggéshez jutunk:

$$x(t) = T \sum_{k \in \mathbb{Z}} \langle X(f), \exp(-j2\pi fkT) \rangle \mathcal{F}^{-1}(\exp(-j2\pi fkT)),$$

ahol

$$\begin{aligned}
\mathcal{F}^{-1}(\exp(-j2\pi fkT)) &= \int_{-W}^{+W} \exp(-j2\pi fkT) \exp(j2\pi ft) df = \\
&= \left[\frac{\exp(j2\pi(t-kT)f)}{j2\pi(t-kT)} \right]_{-W}^{+W} = \frac{1}{T} \frac{\sin(\frac{\pi(t-kT)}{T})}{\frac{\pi(t-kT)}{T}}.
\end{aligned}$$

Felhasználva azt a tényt, hogy az L_2 térben érvényes a Parseval-egyenlőség, az $\{\exp(-j2\pi fkT), \quad k \in \mathbb{Z}\}$ függvények ortogonalitásából közvetlenül következik az $\{\mathcal{F}^{-1}(\exp(-j2\pi fkT)), \quad k \in \mathbb{Z}\}$ függvények ortogonalitása is. Éppen ezért igaz, hogy a

$$\varphi_k(t) = \frac{\sin(\frac{\pi(t-kT)}{T})}{\frac{\pi(t-kT)}{T}} = \text{sinc}_T(t-kT), \quad k \in \mathbb{Z}$$

függvények ortogonális jelkészletet alkotnak, amely jelkészlet bármely sávkorlátozott jelre nézve teljes, mivel bármilyen $(-W, W)$ tartományban sávkorlátozott jel leírható az

$$x(t) = \frac{1}{T} \sum_{k \in \mathbb{Z}} \langle x(t), \text{sinc}_T(t-kT) \rangle \text{sinc}_T(t-kT)$$

alakban, és emellett a Parseval-egyenlőség alapján tudjuk azt is, hogy

$$\|\text{sinc}_T(t-kT)\|^2 = T,$$

mivel

$$\begin{aligned}
\mathcal{F}(\text{sinc}_T(t-kT)) &= T \exp(-j2\pi fkT), \\
\|\mathcal{F}(\text{sinc}_T(t-kT))\|^2 &= T^2 \|\exp(-j2\pi fkT)\|^2,
\end{aligned}$$

és

$$\|\exp(-j2\pi fkT)\|^2 = \frac{1}{T}.$$

Az eddigi eredményeket felhasználva az $x(t)$ függvény $t = jT$ helyen felvett $x(jT)$ mintája az

$$x(jT) = \frac{1}{T} \sum_{k \in \mathbb{Z}} \langle x(t), \text{sinc}_T(t-kT) \rangle \text{sinc}_T(jT-kT) = \frac{1}{T} \langle x(t), \text{sinc}_T(t-jT) \rangle$$

formában adható meg, mivel a $\text{sinc}_T(jT - kT)$ függvény értéke csak a $j = k$ helyen egységnyi, minden más esetben 0, amiből

$$\langle x(t), \text{sinc}_T(t - jT) \rangle = Tx(jT).$$

Visszahelyettesítve ezt az eredményt a korábbi kifejezésbe az $x(t)$ függvény felírható az

$$x(t) = \frac{1}{T} \sum_{k \in \mathbb{Z}} \langle x(t), \text{sinc}_T(t - kT) \rangle \text{sinc}_T(t - kT) = \frac{1}{T} \sum_{k \in \mathbb{Z}} Tx(kT) \text{sinc}_T(t - kT) = \sum_{k \in \mathbb{Z}} x(kT) \text{sinc}_T(t - kT).$$

alakban, azaz a függvény leírható a mintái segítségével.

Ez a **sávhatárolt jelek jól ismert mintavételi tétele**, ami kimondja, hogy egy W sávra sávhatárolt jel a $T = 1/2W$ időnként vett mintái segítségével pontosan visszaállítható.

A korábbi eredmények alapján kimondhatjuk, hogy a

$$\varphi_k(t) = \frac{1}{\sqrt{T}} \frac{\sin\left(\frac{\pi(t-kT)}{T}\right)}{\frac{\pi(t-kT)}{T}} = \frac{1}{\sqrt{T}} \text{sinc}_T(t - kT), \quad k \in \mathbb{Z}$$

jelek minden L_2 térbeli, $W = 1/2T$ sávra sávhatárolt $x(t)$ jel számára teljes ortonormált bázist alkotnak, és az $x(t)$ vetülete a $\varphi_k(t)$ függvényre, azaz az \mathbf{x} vektor k -dik komponense az

$$x_k = \frac{1}{\sqrt{T}} \langle x(t), \text{sinc}_T(t - kT) \rangle = \sqrt{T} x(kT)$$

kifejezés segítségével határozható meg. Mivel a skalárszorzat az ortonormált sorfejtésre nézve invariáns, fennáll az

$$\langle x_h(t), x_l(t) \rangle = \int_{-\infty}^{+\infty} x_h(t) x_l(t) dt = \langle \mathbf{x}_h, \mathbf{x}_l \rangle = T \sum_{k \in \mathbb{Z}} x_h(kT) x_l(kT)$$

egyenlőség.

5.4. Az áteresztő sávi sávkorlátozott jelek leírása a jeltérben

Legyen az $x(t)$ jel sávhatárolt az $[(f_c - \frac{W}{2}, f_c + \frac{W}{2}), f_c \geq \frac{W}{2}]$ sáváteresztő tartományban, azaz legyen a jel Fourier-transzformáltjának a tartója a $(-(f_c + \frac{W}{2}), -(f_c - \frac{W}{2}))$ és az $((f_c - \frac{W}{2}), (f_c + \frac{W}{2}))$ intervallum. Az áteresztő sávi jel sávszélességét azért választottuk ekkorára, hogy az 5.3 fejezetben tárgyalt alapsávi, és az itt tárgyalt áteresztő sávi jelek sávszélessége egyforma legyen (W egyoldalas és $2W$ kétoldalas sávszélesség). Ilyenkor a kétféle rendszerbe jutó Gauss-zaj eredő teljesítménye azonos, ezért a kétféle rendszer jól összehasonlítható.

Az alábbiakban kimutatjuk, hogy, ha a $\{\varphi_k(t), k \in \mathbb{Z}\}$ függvények ortogonális halmazt képeznek az $L_2[0, \frac{W}{2})$ térben, akkor az

$$\{\varphi_k(t) \cos(2\pi f_c t), \varphi_k(t) \sin(2\pi f_c t), \quad k \in \mathbb{Z}\}$$

jelek is ortogonális halmazt alkotnak az $L_2((f_c - \frac{W}{2}), (f_c + \frac{W}{2}))$ térben. Ez az állítás egyszerűen belátható az utóbbi függvények Fourier-transzformáltjainak a felhasználásával. Tudjuk ugyanis, hogy

$$\mathcal{F} [\varphi_k(t) \cos(2\pi f_c t)] = \frac{1}{2} \mathcal{F} [\varphi_k(t) (\exp(j2\pi f_c t) + \exp(-j2\pi f_c t))] = \frac{1}{2} [\Phi_k(f - f_c) + \Phi_k(f + f_c)]$$

és

$$\mathcal{F} [\varphi_k(t) \sin(2\pi f_c t)] = \frac{1}{2j} \mathcal{F} [\varphi_k(t) (\exp(j2\pi f_c t) - \exp(-j2\pi f_c t))] = \frac{1}{2j} [\Phi_k(f - f_c) - \Phi_k(f + f_c)],$$

ahol $\Phi_k(f)$ a $\varphi_k(t)$ függvény Fourier-transzformáltja.

Ezek után vizsgáljuk meg a k -dik és l -dik koszinuszt tartalmazó függvény $(\varphi_k(t) \cos(2\pi f_c t))$ és $\varphi_l(t) \cos(2\pi f_c t)$ skalárszorzatát, azaz számítsuk ki az

$$\langle \varphi_k(t) \cos(2\pi f_c t), \varphi_l(t) \cos(2\pi f_c t) \rangle = \int_{-\infty}^{\infty} \varphi_k(t) \cos(2\pi f_c t) \varphi_l(t) \cos(2\pi f_c t) dt$$

integrált a Parseval-tétel alkalmazásával, miszerint a fenti integrál kiszámítható a függvények Fourier-transzformáltjainak a segítségével is:

$$\langle \varphi_k(t) \cos(2\pi f_c t), \varphi_l(t) \cos(2\pi f_c t) \rangle = \frac{1}{4} \int_{-\infty}^{\infty} [\Phi_k(f - f_c) + \Phi_k(f + f_c)][\Phi_l^*(f - f_c) + \Phi_l^*(f + f_c)] df.$$

Mivel tudjuk, hogy a $\{\varphi_k(t)\}$ függvények sávhatároltak, és $f_c > \frac{W}{2}$, így biztosan igaz, hogy a

$$\Phi_k(f - f_c) \Phi_l^*(f + f_c)$$

és a

$$\Phi_k(f + f_c) \Phi_l^*(f - f_c)$$

szorzat minden frekvencián nulla értékű, ezért ezeket az integrál kiértékelésénél elhagyhatjuk. Ennek alapján a skalárszorzat az

$$\begin{aligned} \langle \varphi_k(t) \cos(2\pi f_c t), \varphi_l(t) \cos(2\pi f_c t) \rangle &= \frac{1}{4} \int_{-\infty}^{\infty} \Phi_k(f - f_c) \Phi_l^*(f - f_c) df + \frac{1}{4} \int_{-\infty}^{\infty} \Phi_k(f + f_c) \Phi_l^*(f + f_c) df = \\ &= \frac{1}{2} \int_{-\infty}^{\infty} \varphi_k(t) \varphi_l(t) dt = \begin{cases} \frac{1}{2}, & k = l \\ 0, & k \neq l \end{cases} \end{aligned}$$

egyenlet segítségével határozható meg, amiből jól látszik, hogy a függvények ortogonálisak egymásra. A fenti számításnál kihasználtuk azt az összefüggést, hogy:

$$\int_{-\infty}^{\infty} \Phi_k(f \mp f_c) \Phi_l^*(f \mp f_c) df = \int_{-\infty}^{\infty} \Phi_k(f) \Phi_l^*(f) df = \int_{-\infty}^{\infty} \varphi_k(t) \varphi_l(t) dt.$$

Hasonló módon elemezzük a k -dik koszinuszt és l -dik szinuszt tartalmazó függvény $(\varphi_k(t) \cos(2\pi f_c t))$ és $\varphi_l(t) \sin(2\pi f_c t)$ skalárszorzatát, azaz számítsuk ki az

$$\langle \varphi_k(t) \cos(2\pi f_c t), \varphi_l(t) \sin(2\pi f_c t) \rangle = \int_{-\infty}^{\infty} \varphi_k(t) \cos(2\pi f_c t) \varphi_l(t) \sin(2\pi f_c t) dt$$

integrált a Parseval-tétel alkalmazásával:

$$\langle \varphi_k(t) \cos(2\pi f_c t), \varphi_l(t) \sin(2\pi f_c t) \rangle = \frac{1}{4j} \int_{-\infty}^{\infty} [\Phi_k(f - f_c) + \Phi_k(f + f_c)][\Phi_l^*(f - f_c) - \Phi_l^*(f + f_c)] df.$$

A számításnál most is csak az azonos frekvenciatartományba eső tagok szorzatait kell figyelembe venni, ennek alapján a skalárszorzat az

$$\langle \varphi_k(t) \cos(2\pi f_c t), \varphi_l(t) \sin(2\pi f_c t) \rangle = \frac{1}{4j} \int_{-\infty}^{\infty} \Phi_k(f - f_c) \Phi_l^*(f - f_c) df - \frac{1}{4j} \int_{-\infty}^{\infty} \Phi_k(f + f_c) \Phi_l^*(f + f_c) df = 0$$

egyenlet segítségével határozható meg, amiből jól látszik, hogy a függvények ortogonálisak egymásra. A fenti számításnál ismét kihasználtuk azt az összefüggést, hogy:

$$\int_{-\infty}^{\infty} \Phi_k(f \mp f_c) \Phi_l^*(f \mp f_c) df = \int_{-\infty}^{\infty} \varphi_k(t) \varphi_l(t) dt.$$

Az eredmények alapján megállapíthatjuk, hogy:

- Az áteresztő sávi sávhatárolt jelek leírásához a

$$\{\varphi_k(t) \cos(2\pi f_c t), \varphi_k(t) \sin(2\pi f_c t), \quad k \in \mathbb{Z}\}$$

jelek ortogonális bázist, a

$$\{\sqrt{2}\varphi_k(t) \cos(2\pi f_c t), \sqrt{2}\varphi_k(t) \sin(2\pi f_c t), \quad k \in \mathbb{Z}\}$$

jelek pedig ortonormált bázist alkotnak.

- Mivel az 5.3 fejezetből tudjuk, hogy a

$$\varphi_k(t) = \frac{1}{\sqrt{T}} \frac{\sin(\frac{\pi(t-kT)}{T})}{\frac{\pi(t-kT)}{T}} = \frac{1}{\sqrt{T}} \text{sinc}_T(t - kT), \quad k \in \mathbb{Z}$$

jelek minden L_2 térbeli, $B = (0, \frac{W}{2}]$ sávra sávhatárolt jel számára teljes ortonormált bázist alkotnak, ha $T = \frac{1}{W}$, ezért az $[(f_c - \frac{W}{2}, f_c + \frac{W}{2}), f_c \geq \frac{W}{2}]$ áteresztő sávban sávhatárolt L_2 térbeli $x(t)$ jelek számára a

$$\left\{ \sqrt{2}\varphi_k(t) \cos(2\pi f_c t), \sqrt{2}\varphi_k(t) \sin(2\pi f_c t) \right\} = \left\{ \frac{\sqrt{2}}{\sqrt{T}} \text{sinc}_T(t - kT) \cos(2\pi f_c t), \frac{\sqrt{2}}{\sqrt{T}} \text{sinc}_T(t - kT) \sin(2\pi f_c t), \quad k \in \mathbb{Z} \right\}$$

jelek alkotnak teljes ortonormált bázist az $L_2 [(f_c - \frac{W}{2}), (f_c + \frac{W}{2})]$ térben.

- Ennek alapján bármely L_2 térbeli $[(f_c - \frac{W}{2}, f_c + \frac{W}{2}), f_c \geq \frac{W}{2}]$ áteresztő sávban sávhatárolt jel leírható az

$$x(t) = \frac{\sqrt{2}}{\sqrt{T}} \sum_{k \in \mathbb{Z}} [x_{ck} \text{sinc}_T(t - kT) \cos(2\pi f_c t) + x_{sk} \text{sinc}_T(t - kT) \sin(2\pi f_c t)], \quad T = \frac{1}{W}$$

formában, ahol

$$x_{ck} = \left\langle x(t), \frac{\sqrt{2}}{\sqrt{T}} \text{sinc}_T(t - kT) \cos(2\pi f_c t) \right\rangle$$

és

$$x_{sk} = \left\langle x(t), \frac{\sqrt{2}}{\sqrt{T}} \text{sinc}_T(t - kT) \sin(2\pi f_c t) \right\rangle.$$

5.5. A fehér Gauss-zaj leírása az általános jeltérben

Az 1.2. fejezet alapján tudjuk, hogy az $N_0/2$ kétoldalas teljesítménysűrűségű fehér Gauss-zajból a jeltér minden dimenziójába függetlenül $N_0/2$ átlagos energia jut, azaz a $\mathbf{v}(t)$ fehér Gauss-zaj és az ortonormált alapsávi bázisfüggvények

$$\mathbf{v}_k = \langle \mathbf{v}(t), \varphi_k(t) \rangle,$$

illetve az ortonormált áteresztő sávi bázisfüggvények

$$\mathbf{v}_{ck} = \left\langle \mathbf{v}(t), \sqrt{2}\varphi_k(t) \cos(2\pi f_c t) \right\rangle, \quad \mathbf{v}_{sk} = \left\langle \mathbf{v}(t), \sqrt{2}\varphi_k(t) \sin(2\pi f_c t) \right\rangle$$

skalárszorzata független, nulla várható értékű és $N_0/2$ szórásnégyzetű valószínűségi változók halmaza. A zaj tehát az ortonormált bázisra vetítve a $\{\mathbf{v}_{ck}\}$ és $\{\mathbf{v}_{sk}\}$ vektorokkal leírható. Fontos megemlíteni,

hogy most is érvényes az irrelevancia tétel, azaz az, hogy a rendszer leírásakor a zajnak csak azokat a komponenseit kell figyelembe venni, amelyek a hasznos jeleket leíró ortonormált vektortérben helyezkednek el, mert ez a rendszer leírása szempontjából elegendő statisztikát biztosít.

Emellett érdemes megjegyezni, hogy a zaj két általános jelre vett vetületének skaláris szorzata a

$$\begin{aligned} \mathbf{E} \{ \langle \mathbf{v}(t), x(t) \rangle \langle \mathbf{v}(t), y(t) \rangle \} &= \mathbf{E} \left\{ \left[\int_{-\infty}^{+\infty} \mathbf{v}(t)x(t)dt \right] \left[\int_{-\infty}^{+\infty} \mathbf{v}(\rho)y(\rho)d\rho \right] \right\} = \\ \mathbf{E} \left\{ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathbf{v}(t)\mathbf{v}(\rho)x(t)y(\rho)dtd\rho \right\} &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{N_0}{2} \delta(t-\rho)x(t)y(\rho)dtd\rho = \\ \frac{N_0}{2} \int_{-\infty}^{+\infty} x(t)y(t)dt &= \frac{N_0}{2} \langle x(t), y(t) \rangle \end{aligned}$$

kifejezés segítségével határozható meg.

Összefoglalva:

Alapsávi sávhatárolt jelek esetén:

- A vevő bemenetére jutó jel $r(t) = x(t) + \mathbf{v}(t)$, ahol $x(t)$ sávhatárolt és korlátozott teljesítményű ($B = (-W, W)$), $\mathbf{v}(t)$ pedig fehér Gauss zaj $N_0/2$ kétoldalas teljesítménysűrűséggel;
- $\{\varphi_k(t)\}$, $k \in \mathbb{Z}$ egy ortonormált bázis, ahol

$$\left\{ \varphi_k(t) = \frac{1}{\sqrt{T}} \text{sinc}_T(t - kT), \quad k \in \mathbb{Z} \right\}$$

és

$$x(t) = \sum_{k \in \mathbb{Z}} x_k \varphi_k(t), \quad x_k = \langle x(t), \varphi_k(t) \rangle;$$

- Ha $x(t)$ sávhatárolt sztochasztikus folyamat, akkor annak minden realizációja leírható egy $\mathbf{x} = \{x_k\}$, $k \in \mathbb{Z}$ valószínűségi változó vektorral;
- Ennek alapján $x(t)$ amplitúdómodulált ortonormált pulzusokból álló összeg, melynek a normánégyzete az

$$\|x(t)\|^2 = \sum_{k \in \mathbb{Z}} \|x_k \varphi_k(t)\|^2 = \sum_{k \in \mathbb{Z}} |x_k|^2$$

kifejezéssel adható meg;

- Ha az $x(t)$ sávhatárolt jel teljesítménykorlátozott, és a korlát P , akkor $\frac{1}{T} \sum_{k \in \mathbb{Z}} |x_k|^2 \leq P$;
- A szimbólumidő $T = 1/2W$, így a szimbólumsebesség $1/T = 2W$;
- A jelenergia korlátja $PT = P/2W$;
- A fehér Gauss-zaj energiája $N_0/2$ szimbólumonként;
- A jel-zaj viszony

$$SNR = \frac{\frac{P}{2W}}{\frac{N_0}{2}} = \frac{P}{N_0 W};$$

- A spektrális hatékonyság ρ [bit/két dimenzió] vagy ρ [bit/s/Hz], amiből az adatsebesség $R = W\rho$ [bit/s].

Áteresztősávi sávhatárolt jelek esetén:

- A vevő bemenetére jutó jel $r(t) = x(t) + \mathbf{v}(t)$, ahol $x(t)$ áteresztősávi sávhatárolt és korlátozott teljesítményű ($B = (f_c - W/2, f_c + W/2)$, $f_c > W/2$) hasznos jel, $\mathbf{v}(t)$ pedig a fehér Gauss zaj $N_0/2$ kétoldalas teljesítménysűrűséggel;
- $\{\varphi_{ck}(t)$, és $\varphi_{sk}(t)\}$, $k \in \mathbb{Z}$ egy ortonormált bázis, ahol

$$\left\{ \varphi_{ck}(t) = \frac{\sqrt{2}}{\sqrt{T}} \text{sinc}_T(t - kT) \cos(2\pi f_c t), \quad k \in \mathbb{Z} \right\},$$

$$\left\{ \varphi_{sk}(t) = \frac{\sqrt{2}}{\sqrt{T}} \text{sinc}_T(t - kT) \sin(2\pi f_c t), \quad k \in \mathbb{Z} \right\}$$

és

$$x(t) = \sum_{k \in \mathbb{Z}} x_{ck} \varphi_{ck}(t) + \sum_{k \in \mathbb{Z}} x_{sk} \varphi_{sk}(t), \quad x_{ck} = \langle x(t), \varphi_{ck}(t) \rangle, \quad x_{sk} = \langle x(t), \varphi_{sk}(t) \rangle;$$

- Ha az $x(t)$ áteresztő sávi sávhatárolt sztochasztikus folyamat, akkor annak minden realizációja leírható egy $\mathbf{x} = \{x_{ck}, x_{sk}\}$, $k \in \mathbb{Z}$ valószínűségi változó vektorral;
- Ennek alapján $x(t)$ amplitúdómodulált ortonormált pulzusokból álló összeg, melynek a normanégyzete az

$$\|x(t)\|^2 = \sum_{k \in \mathbb{Z}} (|x_{ck}|^2 + |x_{sk}|^2)$$

kifejezéssel adható meg:

- Ha az $x(t)$ áteresztő sávi sávhatárolt jel teljesítménykorlátozott, és a korlát P , akkor $\frac{1}{T} \sum_{k \in \mathbb{Z}} (|x_{ck}|^2 + |x_{sk}|^2) \leq P$;
- A szimbólumidő $T = 1/W$, de minden szimbólumidőben két szimbólumot viszünk át, így a valószínűségi szimbólumsebesség $2W = 2/T$;
- A jelenergia korlátja $PT = P/2W$;
- A fehér Gauss-zaj energiája $N_0/2$ szimbólumonként;
- A jel-zaj viszony

$$SNR = \frac{\frac{P}{2W}}{\frac{N_0}{2}} = \frac{P}{N_0 W};$$

- A spektrális hatékonyság ρ [bit/két dimenzió] vagy ρ [bit/s/Hz], amiből az adatsebesség $R = W\rho$ [bit/s].

5.6. Az áteresztő sávi sávhatárolt jelek leírása a komplex számok terében

A fentiekből megállapítható, hogy az alapsávi és az áteresztő sávi modell lényegében ekvivalens egymással. Áteresztő sávi jelek esetén gyakran kényelmes a jelek kvadratúra komponenseihez egy komplex változó valós és képzetes részét hozzárendelni oly módon, hogy alkalmazzuk a szokásos transzformációt a valós számok \mathbb{R}^2 és a komplex számok \mathbb{C} tere között. Eszerint az \mathbb{R}^2 -beli (x, y) két dimenziós vektort megfeleltetjük a \mathbb{C} -beli $x + jy$ komplex számnak. Ez lehetséges, mivel egy független azonos eloszlású nulla várható értékű kétdimenziós v Gauss eloszlású sorozat (dimenzióként $N_0/2$, két dimenzióként, vagyis egy komplex dimenzióban N_0 szórásnégyzettel) megfeleltethető a $v = v_c + jv_s$ sorozatnak, ahol v_c és v_s független, nulla várható értékű Gauss eloszlású valószínűségi változó sorozat komplex dimenzióként $N_0/2$ szórásnégyzettel. Ily módon egy diszkrét idejű komplex modellhez jutunk, ahol $\mathbf{r} = \mathbf{x} + v$ az alábbi jellemzőkkel:

- A szimbólumidő $T = 1/W$, azaz a szimbólumsebesség $W = 1/T$ komplex dimenzióként;
- A v Gauss sorozat nulla várható értékű és N_0 szórásnégyzetű komplex dimenzióként;
- A jel-zaj viszony $SNR = P/WN_0$;
- Az adatsebesség ρ [bit/két dimenzió] vagy ρ [bit/s/Hz].

Eddig ez teljesen megegyezik a korábbiakkal, amikor egy komplex dimenzió helyett két valós dimenziót használtunk. Jegyezzük meg, hogy egy kétdimenziós diszkrét idejű alapsávi modell is átkonvertálható diszkrét idejű komplex modellbe, ha alkalmazzuk a fenti konverziót a valós számok \mathbb{R}^2 és a komplex számok \mathbb{C} tere között.

Fontos azonban hangsúlyozni, hogy a valós számok \mathbb{R}^2 és a komplex számok \mathbb{C} tere közötti konverzió megőrzi egy sor algebrai, geometriai és valószínűségelméleti tulajdonságot, de nem mindegyiket:

- Az additivitás triviálisan megmarad;
- A szorzás nem marad meg (mivel az \mathbb{R}^2 térben két elem szorzása nem egyértelműen definiált, míg a \mathbb{C} térben igen);
- A skaláris szorzat nem marad meg, mivel két ortogonális \mathbb{R}^2 térbeli elemet a \mathbb{C} térben két kollinearitás elemre lehet leképezni;
- A euklideszi norma és az euklideszi távolság megmarad, mivel a kétdimenziós vektor hossza megfelel a komplex szám abszolút értékének;
- Általában, ha v_c és v_s függő valós Gauss eloszlású sorozatok, akkor $v_c + jv_s$ nem megfelelő komplex Gauss eloszlású sorozat, még abban az esetben sem, ha v_c és v_s azonos eloszlásúak;
- Ezzel szemben, ha v_c és v_s független azonos eloszlású nulla várható értékű Gauss eloszlású sorozatok dimenzióként $N_0/2$ szórásnégyzettel, akkor $v_c + jv_s$ egy komplex nulla várható értékű Gauss eloszlású sorozat N_0 szórásnégyzettel komplex dimenzióként.

5.7. Az ortogonális PAM és QAM modulációk

Az ortogonális PAM modulációs rendszerek leírása

Tételezzük fel, hogy az ortogonális PAM jel leírható az

$$x(t) = \sum_k x_k \varphi_k(t)$$

kifejezés segítségével, ahol $\mathbf{x} = \{x_k\}$, $k \in \mathbb{Z}$ az x_k üzenetek valószínűségi változó sorozata és $\{\varphi_k(t) = p(t - kT)\}$, $k \in \mathbb{Z}$ egy ortonormált sorozat, amely egy $p(t)$ elemi jel időbeli eltolásával állítható elő. Ilyenkor általános pulzus amplitúdó modulációról (PAM) beszélünk. Most a jelteret a $\{p(t - kT)\}$, $k \in \mathbb{Z}$ ortonormált sorozat feszíti ki. Az ortogonális PAM rendszer blokkvázlata az 5.1. ábrán látható.

Az elégséges statisztikához most elegendő megfigyelni az

$$r_k = \langle r(t), \varphi_k(t) \rangle = \langle x(t), \varphi_k(t) \rangle + \langle v(t), \varphi_k(t) \rangle = x_k + v_k$$

vetületeket, amiből nyilvánvaló, hogy egyetlen mintából az x_k üzenet visszaállítható, mivel a k -edik kimeneti r_k minta csak a k -edik bemeneti x_k mintától függ. Ilyenkor egyébként szimbólumközi áthallás mentes átvitelről beszélünk. A megfigyeléshez elő kell állítani a fent jelölt skaláris szorzatokat, azaz végre kell hajtani az illesztett szűrést a $p(-t)$ jellel:

$$z(t) = \int_{-\infty}^{+\infty} r(\tau) p(\tau - t) d\tau, \quad z(kT) = \int_{-\infty}^{+\infty} r(\tau) p(\tau - kT) d\tau = r_k.$$

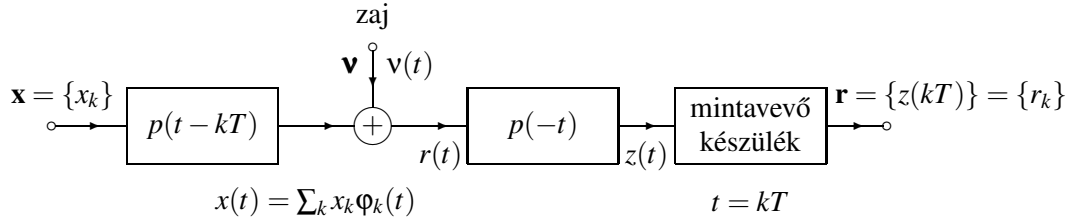
A $\{p(t - kT)\}$, $k \in \mathbb{Z}$ jelek ortonormáltak, ha $p(t)$ teljesíti a Nyquist-feltételt. Ehhez definiáljuk a $g(t) = p(t) * p(-t)$ eredő súlyfüggvényt, amelynek a Fourier-transzformáltja:

$$G(f) = \mathcal{F}(g(t)) = \int_{-\infty}^{+\infty} g(t) e^{-j2\pi ft} dt = |P(f)|^2,$$

ahol $P(f) = \mathcal{F}(p(t))$ a $p(t)$ függvény Fourier-transzformáltja.

Tétel

A Nyquist-kritérium kimondja, ha adott egy $p(t)$ függvény és egy T intervallum, akkor igazak a következő állítások:



5.1. ábra. Az ortogonális PAM modulációs rendszer modellje

- $\{p(t - kT), k \in \mathbb{Z}\}$ függvények ortonormáltak;
- $g(t) = p(t) * p(-t)$ kielégíti a $g(0) = 1$ és $g(kT) = 0$ feltételt minden $k \neq 0$ érték esetén;
- $G(f) = \mathcal{F}(g(t))$, a $g(t)$ Fourier transzformáltja kielégíti a Nyquist-kritériumot, azaz:

$$\frac{1}{T} \sum_{m \in \mathbb{Z}} G\left(f - \frac{m}{T}\right) = 1,$$

minden f értékre.

Bizonyítás

- A két különböző k értékhez tartozó függvény skaláris szorzata a

$$\begin{aligned} \langle p(t - kT), p(t - k'T) \rangle &= \int_{-\infty}^{+\infty} p(t - k'T) p(t - kT) dt \\ &= \int_{-\infty}^{+\infty} p(\tau) p(\tau - kT + k'T) d\tau = g((k - k')T), \end{aligned}$$

mivel

$$g(t) = \int_{-\infty}^{+\infty} p(\tau) p(\tau - t) d\tau;$$

- Tudjuk, hogy a hasznos jel a vevőszűrő kimenetén az

$$x(t) = \sum_{k \in \mathbb{Z}} x_k g(t - kT)$$

alakban írható fel.

Végezzük el ezután az $x(t)$ jel ideális mintavételezését, ami az időtartományban a

$$\sum_{n=-\infty}^{+\infty} \delta(t - nT)$$

függvénnyel való szorzással, a frekvenciatartományban pedig az

$$\mathcal{F} \left(\sum_{n=-\infty}^{+\infty} \delta(t - nT) \right) = \frac{1}{T} \sum_{l=-\infty}^{+\infty} \delta\left(f - \frac{l}{T}\right)$$

függvénnyel képzett konvolúcióval ekvivalens.

Az utóbbi összefüggésnél kihasználtuk, hogy a

$$\sum_{n=-\infty}^{+\infty} \delta(t - nT)$$

periodikus függvény, ezért Fourier-sorba fejthető, és a függvény Fourier-sora a

$$\sum_{n=-\infty}^{+\infty} \delta(t - nT) = \sum_{l=-\infty}^{+\infty} c_l e^{j2\pi \frac{l}{T} t} = \frac{1}{T} \sum_{l=-\infty}^{+\infty} e^{j2\pi \frac{l}{T} t}$$

alakban írható fel, mivel

$$c_l = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \delta(t) e^{-j2\pi \frac{l}{T} t} dt = \frac{1}{T}$$

minden l érték esetén, így

$$\mathcal{F} \left(\sum_{l=-\infty}^{+\infty} c_l e^{j2\pi \frac{l}{T} t} \right) = \frac{1}{T} \sum_{l=-\infty}^{+\infty} \delta \left(f - \frac{l}{T} \right).$$

Számítsuk ki ezután az ideálisan mintavételezett jel két formájának a Fourier-transzformáltját:

a. Tekintsük először az ideális mintavett jel időtartománybeli formájának a Fourier-transzformáltját.

$$\begin{aligned} \mathcal{F} \left(x(t) \sum_{n=-\infty}^{+\infty} \delta(t - nT) \right) &= \int_{-\infty}^{+\infty} \left(\sum_{k=-\infty}^{+\infty} x_k g(t - kT) \right) \left(\sum_{n=-\infty}^{+\infty} \delta(t - nT) \right) e^{-j2\pi f t} dt \\ &= \sum_{n=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} x_k g((n-k)T) e^{-j2\pi f n T} = \sum_{k=-\infty}^{+\infty} x_k e^{-j2\pi f k T}, \end{aligned}$$

ha

$$g((n-k)T) = \begin{cases} 1, & n = k \\ 0, & n \neq k \end{cases}.$$

b. Majd vizsgáljuk meg a jel frekvenciatartománybeli változatának a Fourier-transzformáltját.

$$\begin{aligned} \mathcal{F}(x(t)) * \mathcal{F} \left(\frac{1}{T} \sum_{l=-\infty}^{+\infty} e^{j2\pi \frac{l}{T} t} \right) &= \mathcal{F} \left(\sum_{k=-\infty}^{+\infty} x_k g(t - kT) \right) * \left(\frac{1}{T} \sum_{l=-\infty}^{+\infty} \delta \left(f - \frac{l}{T} \right) \right) \\ &= \left(\sum_{k=-\infty}^{+\infty} x_k G(f) e^{-j2\pi f k T} \right) * \left(\frac{1}{T} \sum_{l=-\infty}^{+\infty} \delta \left(f - \frac{l}{T} \right) \right), \end{aligned}$$

emellett tudjuk, hogy

$$\left(G(f) e^{-j2\pi f k T} \right) * \delta \left(f - \frac{l}{T} \right) = \int_{-\infty}^{+\infty} G(\sigma) e^{-j2\pi \sigma k T} \delta \left(f - \sigma - \frac{l}{T} \right) d\sigma = G \left(f - \frac{l}{T} \right) e^{-j2\pi \left(f - \frac{l}{T} \right) k T},$$

ezért

$$\begin{aligned} \left(\sum_{k=-\infty}^{+\infty} x_k G(f) e^{-j2\pi f k T} \right) * \left(\frac{1}{T} \sum_{l=-\infty}^{+\infty} \delta \left(f - \frac{l}{T} \right) \right) &= \frac{1}{T} \sum_{l=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} x_k G \left(f - \frac{l}{T} \right) e^{-j2\pi \left(f - \frac{l}{T} \right) k T} \\ &= \frac{1}{T} \sum_{l=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} x_k G \left(f - \frac{l}{T} \right) e^{-j2\pi f k T} = \frac{1}{T} \left(\sum_{k=-\infty}^{+\infty} x_k e^{-j2\pi f k T} \right) \left(\sum_{l=-\infty}^{+\infty} G \left(f - \frac{l}{T} \right) \right). \end{aligned}$$

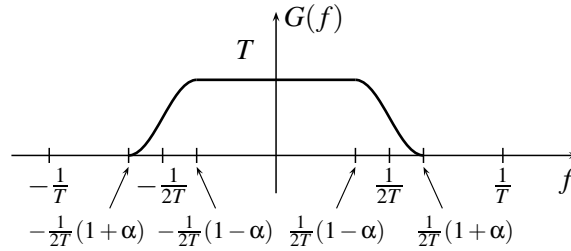
Összehasonlítva az a. és b. pontokban elvégzett számításokat, ortogonális jelek esetén teljesülni kell az

$$\sum_{k=-\infty}^{+\infty} x_k e^{-j2\pi f k T} = \frac{1}{T} \left(\sum_{k=-\infty}^{+\infty} x_k e^{-j2\pi f k T} \right) \left(\sum_{l=-\infty}^{+\infty} G \left(f - \frac{l}{T} \right) \right)$$

egyenlőségnek, ami alapján

$$\frac{1}{T} \sum_{l=-\infty}^{+\infty} G \left(f - \frac{l}{T} \right) = 1,$$

ami a jól ismert Nyquist-feltétel.



5.2. ábra. Az emelt koszinuszos jel spektruma

A Nyquist-feltétel teljesülése esetén nyilvánvaló, hogy amennyiben $p(t)$ egy W sávszélességű alapsávi ekvivalens jel, akkor az ortogonális PAM modulációs rendszer a következő tulajdonságokkal jellemezhető:

- $W = 1/2T$ a PAM rendszer nominális vagy Nyquist sávszélessége, ahol T a szimbólumidő;
- A jelek sávszélessége W nem kevesebb, mint $1/2T$;
- Ha $W = 1/2T$, akkor $G(f) = T$, $-W \leq f \leq W$ és $g(t) = \text{sinc}_T(t)$;
- Ha $1/2T < W \leq 1/T$, akkor $G(f)$ teljesíti a

$$G\left(\frac{1}{2T} + f\right) + G\left(\frac{1}{2T} - f\right) = T, \quad 0 \leq f \leq \frac{1}{2T}$$

feltételt.

Ez utóbbi feltétel teljesülésekor, amit sávszéli Nyquist-szimmetria feltételnek nevezünk, a $G(f) = |P(f)|^2$ lekerekítése szabadon tervezhető az $f > W$ tartományban oly módon, hogy a függvény folytonos legyen és folytonos deriválttal rendelkezzen. Ezt illusztrálja az 5.2. ábra, ahol egy úgynevezett emelt koszinuszos elemi jel $G(f)$ Fourier-transzformáltját tüntettük fel sávszéli Nyquist-szimmetriával.

Ennek alapján megállapíthatjuk, hogy az ortogonális PAM rendszer sávszélessége tetszőlegesen megközelítheti a $W = 1/2T$ értéket, vagy mondhatjuk úgy is, hogy az $1/2T$ sávon kívül eső teljesítmény tetszőlegesen kis értéket vehet fel, anélkül, hogy megsértenénk azt a gyakorlati előírást, hogy a $p(t)$ elemi jel Fourier-transzformáltjának abszolút érték négyzete folytonos legyen és folytonos deriválttal rendelkezzen.

Összefoglalva:

Ortogonalis PAM moduláció esetén:

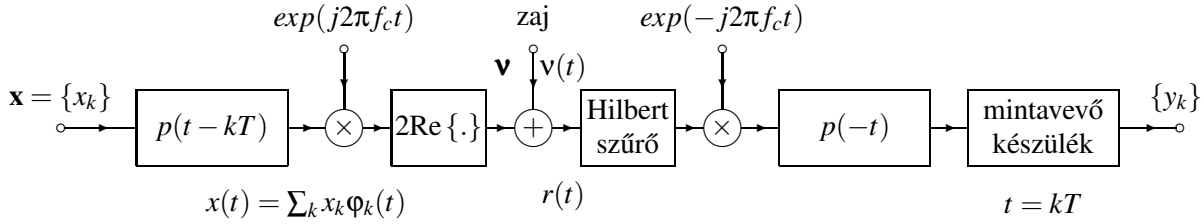
- A szimbólumidő $T = 1/2W$, ahol W a sávszélesség és $2W$ [szimbólum/s] a szimbólumsebesség;
- A jel átlagos energiája $P/2W$ -ben korlátozott;
- A fehér Gauss-zaj minden mintán $N_0/2$ szórásnégyzetű;
- A jel/zaj viszony

$$SNR = \frac{\frac{P}{2W}}{\frac{N_0}{2}} = \frac{P}{N_0 W};$$

- Az adatsebesség ρ [bit/két dimenzió], $R = \rho W$ [bit/s] és a spektrális hatékonyság ρ [bit/s/Hz].

Az ortogonális QAM modulációs rendszerek leírása

Az ortogonális kvadratúra amplitúdó modulált (QAM) rendszer blokkvázlata az 5.3. ábrán látható, ahol a szimbólumidő T , az $\mathbf{x} = \{x_k\}$ és $\mathbf{y} = \{y_k\}$ bemeneti és kimeneti üzenetsorozat elemei komplex véges energiájú valószínűségi változók, amelyek a $p(t - kT)$ elemi jeleket modulálják. Ortonormált esetben a $p(t)$ függvény T idővel eltoló változatai ortonormáltak egymásra. A komplex függvények



5.3. ábra. Az ortogonális QAM modulációs rendszer komplex modellje

terében az $x(t)$ és $y(t)$ függvények skaláris szorzata az

$$\langle x(t), y(t) \rangle = \int_{-\infty}^{+\infty} x(t) y^*(t) dt$$

integrállal adható meg, $p(-t)$ a vevő illesztett szűrőjének súlyfüggvénye, $f_c > 1/2T$ a vivőfrekvencia. A $2\text{Re}\{.\}$ feliratú blokk a bemeneti $f(t)$ jel reális részének a kétszeresét, az $f(t) + f^*(t)$ jelet állítja elő, a komplex Hilbert-szűrő átviteli függvénye pedig 1, ha $f > 0$ és 0, ha $f < 0$.

A rendszerben feltételezzük, hogy $P(f) = 0$ ha $f > f_c$. Ebben az esetben egyszerűen belátható, hogy

- Ha bevezetjük a $g(t) = p(t) * p(-t)$ eredő súlyfüggvényt és az kielégíti a Nyquist-feltételt, akkor az ortogonális QAM rendszer sávszélessége alulról korlátozott a rendszer $W = 1/T$ Nyquist-sávszélességével;
- Ha a vevőszűrő $h(t)$ súlyfüggvénye nem egyezik meg $p(-t)$ -vel, azaz a vevőszűrő nem illesztett szűrő, akkor a kimeneti y_l sorozat a következőképpen számolható:

$$\begin{aligned}
 y_l &= \int_{-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} x_k p(\tau - kT) h(lT - \tau) d\tau = \sum_{k=-\infty}^{+\infty} x_k \int_{-\infty}^{+\infty} p(\tau - kT) h(lT - \tau) d\tau \\
 &= \sum_{k=-\infty}^{+\infty} x_k \int_{-\infty}^{+\infty} p(\tau) h(lT - kT - \tau) d\tau = \sum_{k=-\infty}^{+\infty} x_k z_{l-k},
 \end{aligned}$$

ahol

$$z_{l-k} = \int_{-\infty}^{+\infty} p(\tau) h(lT - kT - \tau) d\tau,$$

ami azt jelenti, hogy a rendszerben lineáris szimbólumközi áthallás keletkezik. Feltételeztük ugyanis, hogy a $g(t) = p(t) * p(-t)$ eredő súlyfüggvény kielégíti a Nyquist-feltételt, de ez általában a $p(t) * h(t)$ függvény esetében nem teljesül, azaz esetünkben $z_{l-k} \neq 0$ ha $l \neq k$;

- A fehér Gauss-zaj minden mintán $N_0/2$ szórásnégyzetű;
- A jel/zaj viszony

$$SNR = \frac{\frac{P}{2W}}{\frac{N_0}{2}} = \frac{P}{N_0 W};$$

- Illesztett szűrő esetén kimeneten megjelenő y_k sorozat felírható az $\mathbf{y} = \mathbf{x} + \mathbf{v}$ formában, ahol \mathbf{v} minden eleme független komplex Gauss-eloszlású valószínűségi változó, és a rendszer jel-zaj viszonya most is P/WN_0 ;
- Egyszerűen belátható, hogy a mintavételi időpont pontatlan meghatározása azaz a mintavétel idejének a $t = lT + \delta$ időpontra történő eltolása szintén lineáris szimbólumközi áthallást okoz. Ez azért következik be, mert még illesztett szűrő esetén is (ha $h(t) = p(-t)$) igaz, hogy a

$$z_{l+\delta-k} = \int_{-\infty}^{+\infty} p(\tau) p(-lT - \delta + kT + \tau) d\tau,$$

sorozat nullától különbözik akkor, ha $l \neq k$, ami azt eredményezi, hogy a k -dik kimeneti y_k minta nem csak a k -dik bemeneti x_k mintától függ;

- Az adó és vevő oldali vivőfrekvenciás jelek ($\exp(j2\pi f_c t)$ és $\exp(-j2\pi f_c t)$) közötti fázistolás az $y(k)$ mintát elforgatja a komplex síkon.

5.8. A kódolatlan rendszerek teljesítőképessége a Shannon-kapacitáshoz viszonyítva

A Shannon féle C csatornakapacitás fehér Gauss-zajos csatorna esetén a

$$C \text{ [bit/két dimenzió]} = W \log_2(1 + SNR)$$

kifejezéssel határozható meg, ahol W a rendszer sávszélessége, SNR pedig a jel-zaj viszony. A csatornakapacitás megadja az R átviteli sebesség felső korlátját, ami felett hibamentesen nem lehet az adott csatornában kommunikálni, azaz

$$R \text{ [bit/két dimenzió]} < C \text{ [bit/két dimenzió]} = W \log_2(1 + SNR).$$

A csatornakapacitás egyúttal felső korlátot ad a ρ [bit/két dimenzió] = ρ [bit/s/Hz] spektrális hatékonyságra is, miszerint

$$\rho < \frac{C}{W} = \log_2(1 + SNR).$$

Ugyanakkor a csatornakapacitás elmélete nem ad olyan konstruktív kódolási eljárást, amivel ezt a korlátot a gyakorlatban meg lehet közelíteni. Ilyen kódolási módszerek kidolgozása a kódoláselmélet alapvető problémája az elmúlt fél évszázadban.

Ebben a fejezetben azt vizsgáljuk, hogy a különböző gyakorlati rendszerek milyen módon képesek a Shannon-kapacitást megközelíteni. A vizsgálat során élesen megkülönböztetjük a teljesítménykorlátozott és a sávkorlátozott rendszereket, mivel az előbbiekben a nominális spektrális hatékonyság kicsi, az utóbbiakban viszont nagy. Teljesítménykorlátozott esetben a rendszerek teljesítőképességét a kódolatlan 2-PAM, míg sávkorlátozott esetben az M -PAM (vagy az ekvivalens $M \times M$ -QAM) teljesítőképességéhez viszonyítjuk. A vizsgálatokat ebben a fejezetben kódolatlan esetekre végezzük el, és ezzel készítjük elő a kódolt rendszerek által biztosított "kódolási nyereség" meghatározását.

A diszkrét idejű fehér Gauss-zajos csatorna modellje

Az előző fejezetben láttuk, hogy egy ortonormált bázis segítségével a folytonos idejű, fehér Gauss-zajos csatornában működő ortogonális PAM és QAM rendszerek univerzálisan leírhatók valós vagy komplex diszkrét idejű vektorokkal, azaz a vevő bemenetére jutó jel leírható az $\mathbf{r} = \mathbf{x} + \mathbf{v}$ egyenlet segítségével, ahol \mathbf{x} a véletlen bemeneti pontsorozat (az üzenetek) vektora, \mathbf{v} pedig a független, nulla várható értékű, azonos eloszlású Gauss-zaj sorozat vektora, melynek minden valós dimenzió irányába $N_0/2$ szórásnégyzetű a vetülete. A korábbiakban azt is láttuk, hogy nincsen lényeges különbség az említett rendszerek valós és komplex diszkrét idejű modelljei között, ezért mostantól kezdve csak a valós modellel fogunk foglalkozni. Visszaidézve a kapcsolatot a diszkrét idejű és a folytonos idejű modellek között, ha a szimbólumsebesség $1/T$ [valódi szimbólum/s]-ban mérve, a spektrális hatékonyság ρ [bit/két dimenzió] és E_s az átlagos szimbólumenergia két dimenzióként, akkor a következő megállapítások érvényesek:

- A rendszer névleges sávszélessége $W = 1/2T$ [Hz];
- A rendszer bitebessége $R = \rho W$ [bit/s], ahol a névleges spektrális hatékonyság ρ [bit/s/Hz];
- Az átlagos jelteljesítmény (a másodpercenkénti átlagos energia) $P = E_s W$;

- A rendszer jel-zaj viszonya $SNR = E_s/N_0 = P/N_0W$;
- A csatorna kapacitása [bit/s]-ban $C = W \log_2(1 + SNR)$.

A normalizált jel-zaj viszony és az E_b/N_0

Ebben a fejezetben bevezetünk két normalizált jel-zaj viszony mértéket, ami a $\rho \leq \log_2(1 + SNR)$ korláthoz kapcsolódik, és amit ezután Shannon-korlátnak fogunk nevezni.

A Shannon-korlát ekvivalens megfogalmazása szerint tetszőleges ρ [bit/két dimenzió] adatsebességű kódolási rendszerben tetszőlegesen kis hibaarányhoz az SNR jel-zaj viszonyának teljesíteni kell az

$$SNR > 2^\rho - 1$$

egyenlőtlenséget. Ebből a normalizált jel-zaj viszony természetesen az

$$SNR_{norm} = \frac{SNR}{2^\rho - 1}$$

értékkel határozható meg, ahol az SNR_{norm} az aktuális SNR -nek a ρ -hoz szükséges minimális SNR -hez viszonyított értéke. A Shannon-korláthoz ezek alapján az $SNR_{norm} = 1$ [0 dB] tartozik. Sőt az SNR_{norm} aktuális értéke [dB]-ben megmutatja azt is, hogy a rendszer a Shannon-korláthoz viszonyítva milyen "munkapontban" működik.

A másik gyakran alkalmazott normalizált jel-zaj viszony érték az E_b/N_0 , ahol E_b az egy bithez tartozó jelenergia, és N_0 a fehér Gauss-zaj egyoldalas teljesítménysűrűsége (a zaj eredő szórásnégyzete két dimenzióként). Érdeemes megjegyezni, hogy $E_b = E_s/\rho$, ahol E_s az átlagos jelenergia két dimenzióként, azaz $E_b/N_0 = E_s/\rho N_0 = SNR/\rho$. Az E_b/N_0 értéket az "egy információs bitre jutó jel-zaj viszonyának" is nevezzük. Az E_b/N_0 értékét általában [dB]-ben adjuk meg.

A fentiek alapján az $SNR > 2^\rho - 1$ Shannon-korlát az E_b/N_0 -ra vonatkoztatva az

$$E_b/N_0 > \frac{2^\rho - 1}{\rho}$$

alakban adható meg.

Jegyezzük meg, hogy az E_b/N_0 alsó korlátja a ρ monoton függvénye. Például $\rho = 2$ -nél $E_b/N_0 > 3/2$ (1.76dB), $\rho = 1$ -nél $E_b/N_0 > 1$ (0dB) és $\rho \rightarrow 0$ esetén $E_b/N_0 > \ln 2$ (-1.59dB), amit az E_b/N_0 abszolút Shannon-korlátjának nevezünk, ugyanis ha $\rho \rightarrow 0$, akkor

$$E_b/N_0 > \lim_{\rho \rightarrow 0} \frac{2^\rho - 1}{\rho} = \lim_{\rho \rightarrow 0} \frac{1 + \rho \ln(2) - 1}{\rho} = \ln(2).$$

Teljesítménykorlátozott és sávkorlátozott rendszerek

Az ideális fehér Gauss-zajos csatornákat kétféle osztályba lehet sorolni, attól függően, hogy nagy vagy kis spektrális hatékonysággal képesek működni. Ennek alapján teljesítménykorlátozott és sávkorlátozott rendszerekről beszélhetünk. A két osztály között nincsen éles határ, de mi a továbbiakban a $\rho = 2$ [bit/két dimenzió] vagy [bit/s/Hz] értéket tekintjük a határnak, annak alapján, hogy éppen ezt a határt lehet elérni egy bináris átviteli rendszerben.

Megjegyezzük, hogy a Shannon-korlát pontos meghatározása a két osztály esetén jelentősen eltér egymástól.

Teljesítménykorlátozott esetben, ha SNR elegendően kicsi, akkor érvényesek a

$$\rho < \log_2(1 + SNR) \approx SNR \log_2 e,$$

és

$$SNR_{norm} \approx \frac{SNR}{\rho \ln 2} = \frac{E_b}{N_0} \log_2 e$$

összefüggések.

Mindez annyit jelent, hogy teljesítménykorlátozott esetben a csatorna kapacitása, illetve az elérhető spektrális hatékonyság egyenesen arányos az SNR jel-zaj viszonyal, és ha $\rho \rightarrow 0$, akkor az SNR_{norm} egy $\log_2 e$ skálafaktor erejéig megegyezik az E_b/N_0 -al. Emellett, ha $\rho \rightarrow 0$, akkor az $SNR_{norm} > 1$ Shannon-korlát megfelel az $E_b/N_0 > \ln 2$ abszolút Shannon-korlátnak.

Sávkorlátozott esetben (ha az SNR elegendően nagy), a csatorna kapacitása, illetve az elérhető spektrális hatékonyság a

$$\rho < \log_2(1 + SNR) \approx \log_2 SNR,$$

egyenlet szerint az SNR logaritmusával arányos, ami alapvetően különbözik a teljesítménykorlátozott esetben érvényes lineáris függéstől. Emellett igaz, hogy

$$SNR_{norm} = \frac{SNR}{2^{\rho} - 1} \approx \frac{SNR}{2^{\rho}}.$$

Mindezek alapján nyilvánvaló, hogy míg teljesítménykorlátozott esetben az SNR megkétszerezése (3 dB-es növelése) a kapacitást megkétszerezi, addig sávkorlátozott esetben a kapacitás (az elérhető spektrális hatékonyság) ugyanilyen feltételek esetén csak 1 [bit/két dimenzió] vagy 1 [bit/s/Hz] értékkel nő.

Nézzünk meg ezután egy-egy tipikus példát a teljesítménykorlátozott és sávkorlátozott rendszerekre.

1. Példa

A szabványos telefon hangcsatorna tipikusan sávkorlátozott rendszer, amit egy ideális sávkorlátozott fehér Gauss-zajos csatornával lehet modellezni. A csatorna sáv szélessége $W \approx 3.5$ kHz, az $SNR \approx 37$ dB. Ebből a Shannon korlát $\rho < \log_2(1 + 10^{3.7}) \approx 37/3$ bit/s/Hz, és $R < 43,000$ bit/s. Az SNR 3 dB-es növelésével az elérhető ρ spektrális hatékonyság csak 1 bit/s/Hz-el, az R átviteli sebesség pedig csak 3500 bit/s-mal nő.

2. Példa

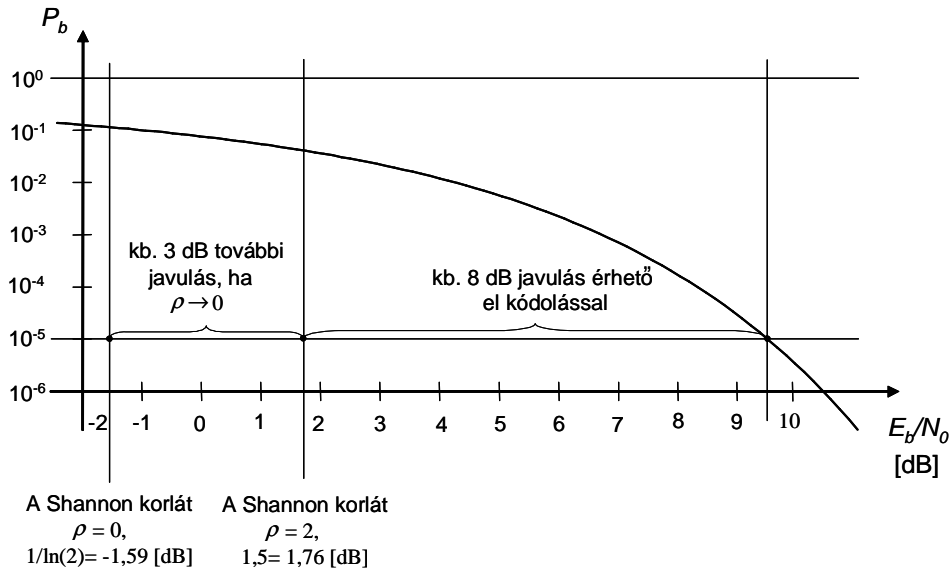
Ezzel szemben az úrtávközlési kommunikációs csatornáknak nincsen sávkorlátozás. Ezért akkora sáv szélességet használhatunk, amekkorára szükségünk van, és a kommunikációt elsősorban a rendelkezésre álló teljesítmény korlátozza. Ebben az esetben az átviteli sebesség felső határát az abszolút Shannon-korlát határozza meg, miszerint $E_b/N_0 > \ln 2 (-1.59 \text{ dB})$. Mivel $E_b/N_0 = P/RN_0$, a Shannon-korlát az $R < (P/N_0)/(\ln 2)$ értékűre adódik. 3 dB-lel növelve a P/N_0 értékét az elérhető átviteli sebesség R bit/s-ban kétszeresére nő.

Megállapítható, hogy a teljesítménykorlátozott és a sávkorlátozott rendszerek alapvetően különböznek egymástól. Teljesítménykorlátozott rendszerekben bináris modulációt és kódolást használhatunk, míg sávkorlátozott rendszerekben a moduláció és kódolás tipikusan nem bináris, hanem többszintű. A teljesítménykorlátozott rendszerekben célszerű minden mennyiséget "egy információs bitre" normalizálni, és az E_b/N_0 a jel-zaj viszony megfelelően normalizált mértéke. Ezzel szemben sávkorlátozott rendszerek esetén sokkal célszerűbb minden mennyiséget "két dimenzióra" normalizálni, és az SNR_{norm} a jel-zaj viszony megfelelően normalizált mértéke.

Ennek alapján a tervezés első lépéseként azt kell rögzíteni, hogy rendszerünk melyik osztályba tartozik, és csak azután lehet megfelelően továbblépni.

5.9. Az M -PAM és az $(M \times M)$ -QAM teljesítőképesége

Ebben a fejezetben meghatározzuk a legegyszerűbb rendszerek, a kódolatlan M -PAM és az $(M \times M)$ -QAM teljesítőképeségét. Ezt a későbbiekben viszonyítási alapnak tekintjük majd. Az alaprendszerek



5.4. ábra. A kódolatlan bináris PAM P_b hibaaránya az E_b/N_0 függvényében

teljesítőképessége és a Shannon-korlát közötti különbség megadja a legbonyolultabb kódolással elérhető maximális nyereséget. Ez meghatározza azt a "játékeret", ami a kódolás számára rendelkezésre áll.

Kódolatlan PAM rendszer

Először foglalkozunk a bináris 2-PAM rendszer tulajdonságaival. A 2-PAM konstellációs diagramja $\mathcal{A} = \{-\alpha, +\alpha\}$, ahol $\alpha > 0$ a skálafaktor, amit úgy választunk meg, hogy az egy bitre jutó átlagos energia értéke $E_b = \alpha^2$ legyen. Ebben a konstellációban a bitsebesség (a nominális spektrális hatékonyság) $\rho = 2$ bit/két dimenzió, és az átlagos bitenergia $E_b = \alpha^2$. A szokásos optimális illesztett szűrős vevő esetén, ha teljesül a szimbólumközi áthallás mentesség feltétele, a detektor a vevőszűrő kimenetén lévő $\{y_k\}$ sorozat alapján az egyes szimbólumokra függetlenül dönt, és döntés eredménye az $\{y_k\}$ sorozat elemeinek előjele, azaz $\hat{x}_k = \text{sgn}(y_k)$. A bithibaarány nyilvánvalóan független attól, hogy az adó éppen melyik jelet küldte, és azonos annak az eseménynek a valószínűségével, hogy az $\sigma^2 = N_0/2$ szórásnégyzetű Gauss-zaj meghaladja az α értéket, azaz $P_b = Q(\alpha/\sigma)$, ahol $Q(\cdot)$ a Gauss-eloszláshoz tartozó hibaintegrál függvény:

$$Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{y^2}{2\sigma^2}\right) dy.$$

Behelyettesítve az $E_b = \alpha^2$ bitenergia és a $\sigma^2 = N_0/2$ szórásnégyzet értéket, a hibaarányra a

$$P_b = Q\left(\sqrt{2\frac{E_b}{N_0}}\right)$$

értéket kapjuk. Az 5.4. ábrán megadtuk a kódolatlan 2-PAM rendszer bithiba arányát az E_b/N_0 függvényében, és bejelöltük a Shannon-korlátot is.

A teljesítménykorlátozott eset és a Shannon-korlát

Teljesítménykorlátozott esetben a kódolatlan bináris pulzus amplitúdó modulációs rendszer (2-PAM) teljesítőképességét tekintjük viszonyítási alapnak, és 2-PAM esetén tudjuk, hogy $\rho = 2$ [bit/s/Hz].

Összehasonlítva a kódolatlan rendszer teljesítőképességét a Shannon-korláttal, meg tudjuk határozni azt a maximálisan elérhető nyereséget, amit tetszőlegesen bonyolult kódolással meg tudunk közelíteni. Teljesítménykorlátozott esetben a normalizált jel-zaj viszony legyen az E_b/N_0 , de ugyanilyen módon használhatjuk az SNR_{norm} értéket is. Mivel $\rho = 2$, $E_b/N_0 = SNR/\rho$ és $SNR_{norm} = SNR/(2^p - 1)$, így $2E_b/N_0 = 3SNR_{norm}$, tehát a viszonyítási alapot jelentő függvény két ekvivalens formában is felírható: $P_b = Q(\sqrt{2E_b/N_0}) = Q(\sqrt{3SNR_{norm}})$.

Az 5.4. ábra általános tervezési lehetőséget biztosít számunkra, mivel, ha például el akarjuk érni a $P_b \approx 10^{-5}$ értéket kódolatlan 2-PAM rendszerrel, akkor tudjuk, hogy $E_b/N_0 \approx 9.6$ dB jel-zaj viszonyra van szükségünk. Ugyanakkor az ábra segítségével a kódolatlan 2-PAM rendszer teljesítőképességét össze tudjuk vetni a Shannon-korláttal is. A 2-PAM rendszer bitsebessége (spektrális hatékonysága) $\rho = 2$ [bit/két dimenzió], amihez az $E_b/N_0 > (2^p - 1)/\rho = 3/2$ (1,76 dB) jel-zaj viszonyra van szükség. Ezért, ha a célunk a $P_b \approx 10^{-5}$ bithibaarány, akkor a kódolatlan 2-PAM-hoz képest kb. 8 dB kódolási nyereséget kell elérni valamilyen hatékony kód alkalmazásával, amelyik szintén biztosítja a $\rho = 2$ bit/két dimenzió adatsebességet.

Sőt, ha nincsen megkötés az alkalmazott sávzélességre, azaz nincs előírás a spektrális hatékonyságra, akkor $\rho \rightarrow 0$ is megengedett, vagyis elérhető az abszolút Shannon-korlát az $E_b/N_0 > \ln 2$ (-1.59 dB) jel-zaj viszonynál. Ha célunk ismét a $P_b \approx 10^{-5}$ bithibaarány biztosítása, akkor Shannon szerint a kódolatlan 2-PAM-hoz képest kb. 11 dB kódolási nyereséget lehet elérni valamilyen hatékony kód alkalmazásával, úgy, hogy megengedjük a zérus értékű spektrális hatékonyságot is.

Kódolatlan M -PAM és $M \times M$ -QAM

Ebben a fejezetben az általánosabb $\mathcal{A} = \{\alpha(\pm 1, \pm 3, \dots, \pm(M-1))\}$ M -PAM konstellációt vizsgáljuk, ahol $\alpha > 0$ ismét a skálafaktor, amit úgy választunk meg, hogy teljesítsük az egy bitre jutó átlagos energia értékére vonatkozó korlátot. A bitsebesség (nominális spektrális hatékonyság) $\rho = 2\log_2 M$ [bit/két dimenzió]. Az egy M -PAM szimbólumra jutó átlagos energia pedig

$$E(\mathcal{A}) = \frac{\alpha^2(M^2 - 1)}{3}.$$

Ezt az állítást egyszerűen be lehet látni, ha kiszámítjuk az 1-től $(M-1)$ -ig terjedő páratlan számok négyzetének az átlagértékét, ami a

$$\frac{2}{M} \sum_{n=1}^{\frac{M}{2}} (2n-1)^2 = \frac{M^2 - 1}{3}$$

kifejezéssel adható meg, ahol M páros.

Ennek alapján a két dimenzióra jutó energia átlagos értéke $E_s = 2E(\mathcal{A}) = 2\alpha^2(M^2 - 1)/3$.

A rendszerben az optimális detektor minden vett y_k szimbólum alapján függetlenül hoz döntést az \hat{x}_k bemeneti üzenetre, ha a rendszer eredő átviteli függvénye teljesíti a Nyquist-feltételt. Ebben az esetben, ha a bemeneti jel αz_k (ahol z_k páratlan egész szám), akkor döntési tartomány az $\alpha [z_k - 1, z_k + 1]$ intervallum (egészen a határokig), kivéve a szélső $\alpha (M-1)$ értékeket, ahol a döntési tartomány $\alpha [M-2, \infty)$. A hibaarány minden $M-2$ -nél kisebb abszolút értékű bemeneti jel esetén a kétszerese annak a valószínűségnek, hogy egy nulla várható értékű, $\sigma^2 = N_0/2$ szórásnégyzetű Gauss-eloszlású valószínűségi változó meghaladja az α értéket, azaz $2Q(\alpha/\sigma)$, míg a $(M-2)$ -nél nagyobb abszolút értékű bemeneti jelek esetén éppen $Q(\alpha/\sigma)$. Ennek alapján az M -PAM rendszer átlagos szimbólum hibaaránya a

$$\Pr(e) = \frac{M-2}{M} 2Q\left(\frac{\alpha}{\sigma}\right) + \frac{2}{M} Q\left(\frac{\alpha}{\sigma}\right) = \frac{2(M-1)}{M} Q\left(\frac{\alpha}{\sigma}\right).$$

$M = 2$ esetén természetesen visszakapjuk a 2-PAM modulációra jellemző $Q(\alpha/\sigma)$ értéket, viszont $M \geq 4$ esetén a $2(M-1)/M$ faktor M növelésekor gyorsan közelíti a 2 értéket, így $\Pr(e) \approx 2Q(\alpha/\sigma)$.

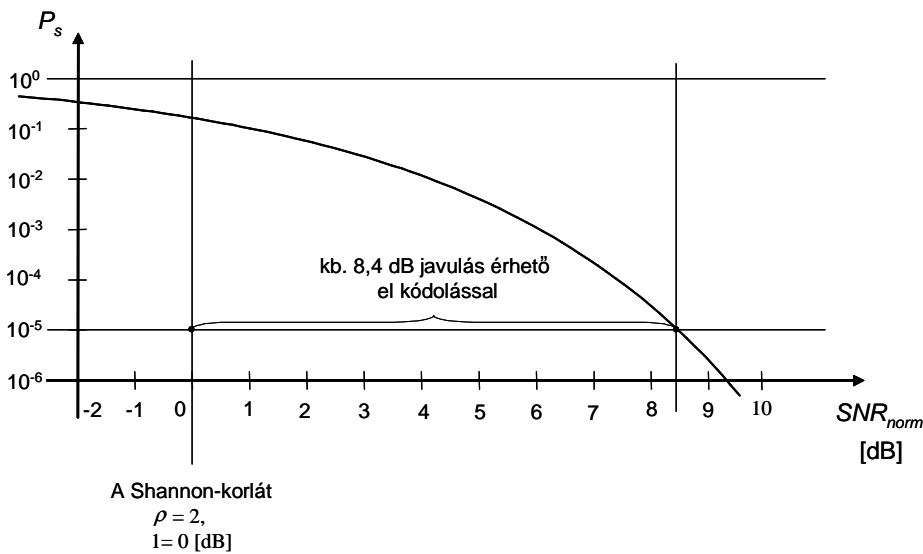
Mivel az $(M \times M)$ -QAM jeltere $\mathcal{A}' = \mathcal{A}^2$ a rendszer ekvivalens két független M -PAM átvitelrel, ezért a fenti a számítást könnyen kiterjeszthetjük az $(M \times M)$ -QAM-re. A bitsebesség (a nominális spektrális hatékonyság) ismét $\rho = 2 \log_2 M$ bit/két dimenzió, és az átlagos két dimenziókénti jelenergia (QAM szimbólumonként) ismét $E_s = 2\alpha^2(M^2 - 1)/3$. A szimbólumonként, és azon belül dimenzióként, történő döntés esetén dimenzióként a hibaarány továbbra is $\Pr(e)$.

A rendszer hibaaránya $(M \times M)$ -QAM szimbólumonként (két dimenzióként) a

$$P_s = 1 - (1 - \Pr(e))^2 = 2\Pr(e) - \Pr(e)^2 \approx 2\Pr(e)$$

kifejezésből számítható. Ezért az $(M \times M)$ -QAM rendszer hibaarányára a $P_s \approx 2\Pr(e) \approx 4Q(\alpha/\sigma)$ értéket kapjuk.

A sávkorlátozott referenciarendszer és a Shannon-korlát



5.5. ábra. A P_s hibaarány kódolatlan $(M \times M)$ -QAM esetén az SNR_{norm} függvényében

Sávkorlátozott esetben a kódolatlan $(M \times M)$ -QAM, $M \geq 4$ rendszer teljesítőképességét tekintjük a továbbiakban viszonyítási alapnak, ezért minden mennyiséget két dimenzióra normalizálunk, és az SNR_{norm} értékkel mérjük a jel-zaj viszonyt.

$M \geq 4$ esetben a két dimenzióra eső hibaarány a

$$P_s \approx 4Q\left(\frac{\alpha}{\sigma}\right)$$

alakban adható meg.

Ide behelyettesítve az $E_s = 2\alpha^2(M^2 - 1)/3$ két dimenziókénti átlagos jelenergiát, a zaj $\sigma^2 = N_0/2$ szórásnégyzetét, az

$$SNR_{norm} = \frac{SNR}{2^{\rho} - 1} = \frac{E_s/N_0}{M^2 - 1}$$

normalizált jel-zaj viszonyt, és $(M^2 - 1)$ -gyel egyszerűsítve a hibaarányra a

$$P_s \approx 4Q(\sqrt{3SNR_{norm}})$$

értéket kapjuk.

Megjegyezzük, hogy ez a görbe nem függ M -től, ami arra utal, hogy az SNR_{norm} jól választott referencia a sávkorlátozott rendszerekben. A sávkorlátozott rendszerek P_s referenciagörbét az SNR_{norm} függvényében az 5.5. ábrán adtuk meg kódolatlan $(M \times M)$ -QAM esetére.

Az 5.5. ábra egy másik univerzális tervezési eszközt jelent számunkra. Ha például kódolatlan $(M \times M)$ -QAM rendszerrel el akarjuk érni a $P_s \approx 10^{-5}$ hibaarányt, akkor tudjuk, hogy $SNR_{norm} \approx 8.4$ dB jel-zaj viszonyra van szükségünk (megjegyezzük, hogy az SNR_{norm} már normalizálva van a spektrális hatékonyságra, szemben az E_b/N_0 -lal).

A Shannon-korlát az SNR_{norm} -ra vonatkoztatva $SNR_{norm} > 1$ (0dB), így Shannon szerint hatékony kódolással még körülbelül 8.4 dB kódolási nyereséget tudunk elérni.

6. fejezet

Kis jelterek teljesítőképessége

Ebben a fejezetben megvizsgáljuk azt, hogy a kis és közepes méretű jelkonstellációk esetén hogyan becsülhető a rendszer teljesítőképessége diszkrét idejű fehér Gauss-zajos csatornában. Maximum likelihood döntési algoritmus esetén (ha az üzenetek eloszlása egyenletes a konstellációs pontok felett) az optimális döntési szabály a minimális távolság alapján fogalmazható meg, mindig arra a szimbólumra döntünk, amelyiknek a vett jeltől mért euklideszi távolsága minimális.

Célunk az, hogy a 2. fejezetben megismert uniós korlát alkalmazásával meghatározzuk a különböző kis és közepes méretű jelkonstellációk teljesítőképességét, és rendszerekhez tartozó kódolási nyereség értékét teljesítménykorlátozott és sávkorlátozott esetben, összehasonlítva ezek paramétereit rendre a 2-PAM és az $(M \times M)$ -QAM rendszerekével.

6.1. Jelkonstellációk fehér Gauss-zajos csatorna esetén

Általában elmondhatjuk, hogy diszkrét idejű fehér Gauss-zajos csatorna esetén egy kódolási rendszer az $\mathbf{y} = \mathbf{x} + \mathbf{v}$ egyenlettel írható le, ahol \mathbf{v} a Gauss-zaj vetületeinek független, nulla várható értékű, azonos eloszlású sorozata. A rendszerben a bemeneti bitsorozatot egy valós \mathbf{x} szimbólumsorozatra képezzük le, amit átviszünk a csatornán (ezt az eljárást kódolásnak nevezzük). A vevőoldalon pedig a vett \mathbf{y} valós szimbólumsorozat megfigyelésével előállítjuk a bemeneti \mathbf{x} bitsorozat $\hat{\mathbf{x}}$ becslését (ezt az eljárást dekódolásnak nevezzük). A dekódolás tehát nem más, mint az \mathbf{y} valós szimbólum sorozat leképzése a kimeneti $\hat{\mathbf{x}}$ bitsorozatra.

A továbbiakban először a Shannon által is használt N hosszúságú blokk kódokkal fogunk foglalkozni, ahol az átvinni kívánt \mathbf{x} sorozatot az $(\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{x}_{k+1}, \dots, \mathbf{x}_M)$, $\mathbf{x}_k \in \mathbb{R}^N$ blokk kód N hosszúságú kódszavai közül kerül ki. Ezután feltételezzük, hogy a blokk kód kódszavainak a száma M , és az egyes kódszavakat ebből az M elemű halmazból függetlenül, egyenletes eloszlással választjuk ki. Meg kell jegyezni, hogy a blokk kód nem az egyetlen kódolási lehetőség, vannak ettől eltérő szerkezetű kódok is, például a konvolúciós kódok vagy a trellis kódok.

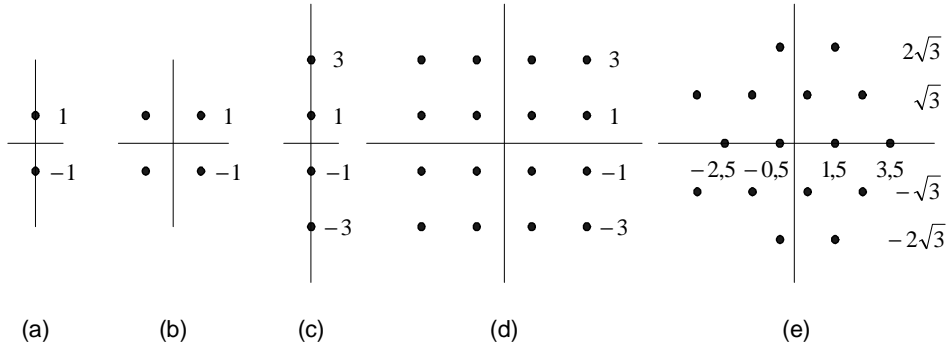
A kódszavak M számát általában 2 valamilyen hatványára választjuk, és a kódszavakat egy $\log_2 M$ hosszúságú bemeneti bitsorozathoz rendeljük. Ha feltételezzük, hogy a bemeneti bitsorozat minden eleme független, azonos valószínűségű bitek sorozata (memóriamentes szimmetrikus forrás, ahol minden bit 0,5 valószínűséggel veszi fel a logikai "1" vagy a logikai "0" értéket), akkor az átvinni kívánt \mathbf{x} sorozat is egyenletes eloszlású lesz a kódszavak $(\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{x}_{k+1}, \dots, \mathbf{x}_M)$, $\mathbf{x}_k \in \mathbb{R}^N$ halmaza felett. A továbbiakban majdnem minden esetben azt fogjuk feltételezni, hogy a \mathbf{x} egyenletes eloszlású, azért, mert ez a legrosszabb eset, ugyanakkor ez a feltétel a valóságos rendszerekben is gyakran teljesül, mivel a forráskódolás hatékonyságának növelése érdekében célszerű megközelíteni a független azonos eloszlású bitsorozat esetét (minimális redundanciájú forráskód).

Digitális kommunikáció esetén általában csak magával a kóddal foglalkozunk, és nem vizsgáljuk azt, hogy az egyes kódszavakat hogyan rendeljük hozzá az üzenetekhez. Más szempontból, csak a

kódszavak elrendezésének van jelentősége.

Ha a blokk kód hossza N és a kódszavak száma M relatíve kicsi, akkor fehér Gauss-zajos csatornában a kódot jelhalmaznak, jel ABC-nek vagy konstellációs elrendezésnek szoktuk nevezni. Ha az N értéke 1 vagy 2 a PAM vagy QAM jelzési tartományában mérve, akkor tipikusan "kódolatlan" rendszerről beszélünk.

A 6.1. ábrán néhány egy- és kétdimenziós jelkonstellációt tüntettünk fel.



6.1. ábra. Kódolatlan jelkonstellációk, (a) 2-PAM; (b) (2×2) -QAM; (c) 4-PAM; (d) (4×4) -QAM; (e) hexagonális 16-QAM

Egy N -dimenziós jelkonstellációt (jelhalmazt, ABC-t) a továbbiakban

$$\mathcal{A} = \{\mathbf{a}_j, 1 \leq j \leq M\}$$

kifejezéssel fogjuk jelölni, melynek az $\mathbf{a}_j \in \mathbb{R}^N$ elemeit jelvektoroknak (jelpontoknak, N hosszúságú sorozatoknak) fogjuk nevezni. Az $\mathcal{A} = \{\mathbf{a}_j, 1 \leq j \leq M\}$ jelkonstelláció alapparaméterei az alábbiak: a jelkonstelláció dimenziója N ; a jelvektorok száma M ; a jeltér átlagos energiája $E(\mathcal{A}) = \frac{1}{M} \sum_j \|\mathbf{a}_j\|^2$ és a minimális euklideszi távolság $d_{\min}(\mathcal{A})$, ami a zajérzékenység elemi mértéke. A jelkonstelláció másodlagos paramétere a legközelebbi szomszédok átlagos száma $K_{\min}(\mathcal{A})$ (a $d_{\min}(\mathcal{A})$ távolságra lévő pontok átlagos száma)

Ezekből a jellemzőkből az alábbi adatok származtathatók:

- A bitsebesség (a névleges spektrális hatékonyság): $\rho = (2/N) \log_2 M$ [bit/két dimenzió];
- A két dimenzióra eső átlagos energia: $E_s = (2/N) E(\mathcal{A})$ vagy $E_b = E(\mathcal{A}) / (\log_2 M) = E_s / \rho$;
- A rendszer normalizált energiájának a mérőszámai, amelyek függetlenek a dimenzió nagyságától: $d_{\min}^2(\mathcal{A}) / E(\mathcal{A})$, $d_{\min}^2(\mathcal{A}) / E_s$ vagy $d_{\min}^2(\mathcal{A}) / E_b$.

Például a 6.1. ábrán megadott elrendezések esetén a bitsebesség a 2-PAM és (2×2) -QAM esetén 2 bit/két dimenzió, míg a többi esetben 4 bit/két dimenzió. A két dimenzióra eső átlagos energia a 2-PAM és (2×2) -QAM esetén $E_s = 2$, a 4-PAM és (4×4) -QAM esetén $E_s = 10$, míg a hexagonális 16-QAM esetén $E_s = 8.75$. A minimális euklideszi távolság négyzete minden esetben $d_{\min}^2 = 4$, a legközelebbi szomszédok átlagos száma pedig rendre $K_{\min}(\mathcal{A}) = 1; 2; 1, 5; 3$ és 4, 125.

A konstellációk Descartes-szorzata

A konstellációk Descartes-szorzatának kérdése azért vetődik fel, mert az $(M \times M)$ -QAM konstelláció két M -PAM konstelláció Descartes szorzata. Általánosan egy \mathcal{A}^K Descartes-szorzat típusú konstelláció nem más, mint az elemi \mathcal{A} konstelláció pontjai összes lehetséges K sorozatának halmaza, azaz

$$\mathcal{A}^K = \{(\mathbf{x}_{j_1}, \mathbf{x}_{j_2}, \dots, \mathbf{x}_{j_K}), j_1, j_2, \dots, j_K \in [1, 2, \dots, M], \mathbf{x}_k \in \mathcal{A}\}.$$

Ha az \mathcal{A} dimenziója és mérete N és M , akkor az \mathcal{A}^K dimenziója $N' = KN$, a mérete $M' = M^K$.

Példa

(a) Ha $\mathcal{A}' = \mathcal{A}^K$, akkor $N' = KN$, $\log_2 M' = \log_2 M^K = K \log_2 M$ és $K_{min}(\mathcal{A}^K) = K K_{min}(\mathcal{A})$, azaz ezek a paraméterek K -szorosára nőnek, viszont az \mathcal{A}^K normalizált paraméterei azonosak az eredeti \mathcal{A} jeltérével. Ez a tény az alábbiakból látható be:

•

$$E'_s = \frac{2}{N'} E(\mathcal{A}^K) = \frac{2K}{KN} E(\mathcal{A}) = E_s,$$

mivel

$$E(\mathcal{A}^K) = \frac{1}{M^K} \sum_{j_1=1}^M \sum_{j_2=1}^M \dots \sum_{j_K=1}^M \left(\|\mathbf{x}_{j_1}\|^2 + \|\mathbf{x}_{j_2}\|^2 + \dots + \|\mathbf{x}_{j_K}\|^2 \right) = K \frac{M^{K-1}}{M^K} \sum_{k=1}^M \|\mathbf{x}_k\|^2 = KE(\mathcal{A});$$

•

$$E'_b = \frac{E(\mathcal{A}^K)}{\log_2 M^K} = \frac{KE(\mathcal{A})}{K \log_2 M} = E_b;$$

• és triviálisan

$$d_{min}^2(\mathcal{A}^K) = d_{min}^2(\mathcal{A}).$$

Az állítások a 6.1. ábra alapján az $(M \times M)$ -QAM esetére egyszerűen igazolhatók.

Emellett érdemes megjegyezni, hogy nincs különbség az \mathcal{A} és az \mathcal{A}^K jelterek \mathbf{x} véletlen bemeneti sorozatai között, azaz nincs különbség az M -PAM és az $(M \times M)$ -QAM véletlen bemeneti sorozatok között. Mindez azért van így, mert a Descartes-szorzat típusú konstellációk esetében feltételeztük a független átvitelt a Descartes-szorzat egyes "tényezői" esetében. Ennek alapján a Descartes-szorzat típusú \mathcal{A}^K konstellációk ekvivalensek az eredeti \mathcal{A} konstellációval (annak egyszerű "változatai"), mivel azonos ρ , E_s , E_b és d_{min}^2 paraméterekkel rendelkeznek.

A későbbiekben az \mathcal{A} térbeli \mathcal{C} kódot úgy definiáljuk, hogy $\mathcal{C} \subset \mathcal{A}^K$ részhalmaza az \mathcal{A}^K Descartes-szorzat típusú konstellációnak. Általában egy \mathcal{A} halmaz feletti \mathcal{C} kód kisebb bitsebességű, mint maga az \mathcal{A} konstelláció, de a d_{min}^2 értéke nagyobb, mint az \mathcal{A} konstellációé, így remélhetjük, hogy az ilyen kód alkalmazásával "kódolási nyereséget" érhetünk el. A következőkben kivétel nélkül ilyen kódokkal fogunk foglalkozni.

Minimális távolságú dekódolás

A dekódolás most is a vett jel leképezését jelenti a küldött információsorozat becslt értékére (néha a demoduláció ennél többet jelent, de induljunk ki ebből a leegyszerűsítő elképzelésből). Ha blokk kódolási eljárást alkalmazunk, akkor a dekódolást is blokkonként kell elvégezni. Egyszerűen belátható, hogy fehér Gauss-zajos csatornában ezzel semmit nem veszítünk az általánosságból. Visszaidézve a blokkból blokkba dekódolás alapelvét, fehér Gauss-zajos csatorna és egyenletes bemeneti üzeneteloszlás esetén, az optimális dekódolás a minimális távolságú (MD) dekódolás.

A blokkból blokkba dekódolás esetén a csatornamodell az

$$\mathbf{y} = \mathbf{x} + \mathbf{v}$$

kifejezéssel adható meg, ahol \mathbf{y} a vett jelek, \mathbf{x} az üzenetszimbólumok és \mathbf{v} a Gauss-zaj N hosszúságú sorozata. Az üzenetsorozatot egyenletes eloszlással választjuk ki az N darab N dimenziós \mathbf{a}_j elemekből álló \mathcal{A} jelkonstellációból. A zaj valószínűségi sűrűségfüggvénye

$$p_{\mathbf{v}}(\mathbf{z}) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left(-\frac{\|\mathbf{z}\|^2}{2\sigma^2}\right),$$

ahol $\sigma^2 = N_0/2$ a zaj szórásnégyzete.

Digitális átvitel esetén általában minimális hibaarányú dekódolásra törekszünk, azaz egy adott \mathbf{y} sorozat ismeretében úgy kell megválasztanunk az $\hat{\mathbf{a}} \in \mathcal{A}$ jelpontot, hogy a $\Pr(e)$ döntési hiba minimális legyen.

Mivel annak a valószínűsége, hogy $\hat{\mathbf{a}}$ becslés helyes, azonos a $p(\hat{\mathbf{a}}|\mathbf{y})$ a posteriori valószínűséggel, a minimális hibaarányú dekódolás azonos a maximum-a-posteriori (MAP) döntéssel, miszerint válasszuk meg $\hat{\mathbf{a}} \in \mathcal{A}$ értékét úgy, hogy a $p(\hat{\mathbf{a}}|\mathbf{y})$ legyen maximális az összes $p(\mathbf{a}_j|\mathbf{y})$, $\mathbf{a}_j \in \mathcal{A}$ között.

A Bayes-szabály kimondja, hogy

$$p(\mathbf{a}_j|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{a}_j)p(\mathbf{a}_j)}{p(\mathbf{y})}.$$

Ha \mathbf{a}_j egyenletes eloszlású, $p(\mathbf{a}_j) = 1/M$ minden j -re, akkor a MAP döntés ekvivalens a maximum likelihood (ML) döntéssel, miszerint válasszuk meg $\hat{\mathbf{a}} \in \mathcal{A}$ értékét úgy, hogy a $p(\mathbf{y}|\hat{\mathbf{a}})$ legyen maximális az összes $p(\mathbf{y}|\mathbf{a}_j)$, $\mathbf{a}_j \in \mathcal{A}$ között.

Felhasználva a zaj valószínűségi sűrűségfüggvényét, azt kapjuk, hogy

$$p(\mathbf{y}|\mathbf{a}_j) = p_v(\mathbf{y} - \mathbf{a}_j) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left(-\frac{\|\mathbf{y} - \mathbf{a}_j\|^2}{2\sigma^2}\right).$$

Ezért az ML döntés ekvivalens a minimális távolságú (MD) döntéssel, azaz válasszuk meg $\hat{\mathbf{a}} \in \mathcal{A}$ értékét úgy, hogy a $\|\mathbf{y} - \hat{\mathbf{a}}\|^2$ legyen minimális az összes $\|\mathbf{y} - \mathbf{a}_j\|^2$, $\mathbf{a}_j \in \mathcal{A}$ esetén.

Összefoglalva: egyenletes eloszlású bemenet és független azonos eloszlású Gauss-zaj esetén a minimális hibaarányú döntési szabály azonos a minimális távolságú döntési szabállyal. Ezért ettől kezdve csak a minimális távolságú döntési szabállyal fogunk foglalkozni, amit igen egyszerűen lehet geometriailag is interpretálni.

Példa

Descartes-szorzat típusú \mathcal{A}^K konstelláció esetén, a minimális távolságú döntési szabályt az $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_K)$ elemekre függetlenül végre lehet hajtani, mivel

$$\min_{\mathbf{a}_j} \|\mathbf{y} - \mathbf{a}_j\|^2 = \sum_{l=1}^K \min_{\mathbf{a}_{j_l}} \|\mathbf{y}_l - \mathbf{a}_{j_l}\|^2.$$

Ha $\Pr(e)$ a minimális távolságú döntés hibaaránya az \mathcal{A} konstelláció esetén, akkor az \mathcal{A}^K konstellációnál a hibaarány a

$$\Pr(e)' = 1 - (1 - \Pr(e))^K$$

kifejezéssel határozható meg, mivel $\Pr(e)$ annak az eseménynek a valószínűsége, hogy az $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_K)$ vett sorozat egyik elemére a minimális távolságú döntési algoritmussal hibás döntést hoztunk, $(1 - \Pr(e))^K$ pedig annak az eseménynek a valószínűsége, hogy a K elem mindegyikére helyes volt a döntésünk, így $1 - (1 - \Pr(e))^K$ valóban a \mathcal{A}^K konstelláció hibavalószínűsége.

Ha $\Pr(e) \ll 1$, akkor

$$1 - (1 - \Pr(e))^K = 1 - \sum_{k=0}^K (-1)^k \binom{K}{k} \Pr(e)^k \approx 1 - (1 - K\Pr(e)) = K\Pr(e).$$

Ha $\mathcal{A} = \{\pm\alpha\}$ és $\mathcal{A}' = \mathcal{A}^K$, akkor ez a rendszer ekvivalens azzal, hogy egymástól függetlenül K számú 2-PAM szimbólumot viszünk át a csatornán. Geometriai értelemben az \mathcal{A}' halmaz elemei egy K dimenziós, 2α élű kocka csúcsaihoz rendelhetők.

Felhasználva a korábbi eredményeket az $\mathcal{A}' = \mathcal{A}^K$, K dimenziós kocka típusú konstelláció esetén a dimenziók száma $N' = K$, a kód mérete $M' = 2^K$, a bitsebesség (a nominális spektrális hatékonyság) $\rho = 2$ bit/két dimenzió, az átlagos energia $E(\mathcal{A}') = K\alpha^2$, a minimális euklideszi távolság

négyzete $d_{\min}^2(\mathcal{A}') = 4\alpha^2$, az átlagos bitenergia $E_b = \alpha^2$, és a legközelebbi szomszédok átlagos száma $K'_{\min}(\mathcal{A}') = K$.

Ennek alapján a $\Pr(e)'$ hibaarány közelítőleg K -szorosa az \mathcal{A} hibaarányának:

$$\Pr(e)' \approx KQ\left(\sqrt{2\frac{E_b}{N_0}}\right),$$

és a bit hibaarány a

$$P_b' \approx \frac{\Pr(e)'}{K} \approx Q\left(\sqrt{2\frac{E_b}{N_0}}\right)$$

kifejezéssel közelíthető.

6.2. A teljesítőképesség vizsgálata a teljesítménykorlátozott tartományban

A teljesítménykorlátozott tartományt korábban úgy definiáltuk, hogy benne a ρ spektrális hatékonyság nem nagyobb, mint 2 bit/két dimenzió. Ebben a tartományban minden értéket célszerű egy bitre normalizálni, így a jel-zaj viszony normalizált értéke E_b/N_0 . Itt a referenciajel a kódolatlan 2-PAM moduláció az $\mathcal{A} = \{\pm\alpha\}$ konstellációval, vagy az ezzel ekvivalens \mathcal{A}^K konstelláció. Ennél a konstellációnál a bitsebesség (a nominális spektrális hatékonyság) $\rho = 2$ bit/két dimenzió, az átlagos bitenergia $E_b = \alpha^2$, a minimális euklideszi távolság négyzete $d_{\min}^2(\mathcal{A}) = 4\alpha^2$ és a minimális távolságban lévő szomszédok átlagos száma $K_b(\mathcal{A}) = 1$. Emellett a korábbiakból tudjuk, hogy az uniós korlát szerint a bithiba a

$$P_b \approx Q\left(\sqrt{2\frac{E_b}{N_0}}\right)$$

értékkel közelíthető.

Ez a referencia diagram az 5.4. ábrán látható. Az \mathcal{A} jeltér effektív kódolási nyeresége $\gamma_{eff}(\mathcal{A})$ egy adott P_b hibaarány esetén azonos a vizsgált jeltérben és a referencia jeltérben (a kódolatlan 2-PAM rendszerben) az ehhez a hibaarányhoz szükséges E_b/N_0 jel-zaj viszony dB-ben kifejezett különbségével.

Például, korábban láttuk, hogy a maximális lehetséges effektív kódolási nyereség $P(b) \approx 10^{-5}$ -nél közelítőleg 11.2 dB. Kisebb $P(b)$ esetén a maximális effektív kódolási nyereség nagyobb, nagyobb $P(b)$ esetén kisebb. E szerint a definíció szerint az effektív kódolási nyereség magában foglalja azt a nyereséget is, ami a $\rho < 2$ [bit/két dimenzió] névleges spektrális hatékonyság alatti kommunikáció miatt keletkezik, ami közelítőleg 3.35 dB. Ha a ρ állandó értékű, például $\rho = 2$ [bit/két dimenzió], akkor a maximális effektív kódolási nyereség $P(b) \approx 10^{-5}$ -nél közelítőleg 8 dB. Ha a ρ (sáv szélesség) értéke korlátozott, akkor a $P(b)$ hibaarányt célszerű az SNR_{norm} függvényében ábrázolni, hogy megmutassuk, hogy az \mathcal{A} jeltér milyen távol van az elérhető kapacitástól.

Az uniós korlát lehetővé teszi, hogy becslést adjunk az effektív kódolási nyereség értékére az alábbi módon. Az egy bitre jutó hibavalószínűség (általában nem azonos a bit hibavalószínűséggel) a

$$P_b = \frac{\Pr(e)}{\log_2 |\mathcal{A}|} \approx \frac{K_{\min}}{\log_2 |\mathcal{A}|} Q\left(\sqrt{\frac{d_{\min}^2(\mathcal{A})}{2N_0}}\right)$$

kifejezéssel határozható meg.

A teljesítménykorlátozott tartományban a névleges kódolási nyereséget a

$$\gamma_c(\mathcal{A}) = \frac{d_{\min}^2(\mathcal{A})}{4E_b}$$

formában határozhatjuk meg. Ezt úgy normalizáltuk, hogy 2-PAM, $\gamma_c(\mathcal{A}) = 1$ legyen. Mivel a nominális kódolási nyereség egy szorzófaktor a $Q(\sqrt{\cdot})$ függvény argumentumában, ezért általában dB-ben mérjük. Az uniós korlát szerint

$$P_b \approx K_b(\mathcal{A}) Q \left(\sqrt{2 \frac{\gamma_c(\mathcal{A}) E_b}{N_0}} \right),$$

ahol

$$K_b(\mathcal{A}) = \frac{K_{min}(\mathcal{A})}{\log_2 |\mathcal{A}|}$$

az egy átvitt bitre jutó legközelebbi szomszédok átlagos száma. Jegyezzük meg, hogy a P_b -re vonatkozó kifejezés 2-PAM esetében pontos.

Az uniós korlát szerint adott $\gamma_c(\mathcal{A})$ és $K_b(\mathcal{A})$ esetén a hibaarány diagramot az alábbi szabály szerint lehet megszerkeszteni: a referencia diagramot balra kell eltolni $\gamma_{eff}(\mathcal{A})$ értékkel dB-ben, és felfelé kell eltolni $K_b(\mathcal{A})$ értékkel szintén dB-ben, mivel a P_b értékét logaritmikus skálán ábrázoltuk. Ez a tény is megerősíti azt, hogy miért célszerű a hibaarányt log-log skálán ábrázolni, ahol a jel-zaj viszony dB-ben van megadva.

Ennek alapján ha $K_b(\mathcal{A}) = 1$, akkor az effektív kódolási nyereség $\rho_{eff}(\mathcal{A})$ azonos a $\rho_c(\mathcal{A})$ nominális kódolási nyereséggel minden P_b értéknél, az uniós korlát pontosságával.

Egyébként, ha $K_b(\mathcal{A}) > 1$, akkor az effektív kódolási nyereség kisebb, mint a nominális kódolási nyereség, egy olyan faktorról, ami a $P_b - E_b/N_0$ görbe meredekségétől függ egy adott P_b értéknél. Ha $P_b \approx 10^{-5}$, akkor egy meglehetősen pontos ökölszabály szerint, nem nagyon nagy $K_b(\mathcal{A})$ értékeknél $K_b(\mathcal{A})$ duplázódása esetén az effektív kódolási nyereség közelítőleg 0,2 dB értékkel csökken, azaz

$$\gamma_{eff}(\mathcal{A}) \approx \gamma_c(\mathcal{A}) - (0,2) \log_2(K_b(\mathcal{A})) \quad \text{dB-ben.}$$

Ennél pontosabb becslés az uniós korlát alapján adható.

Ortogonalis és azzal rokon jeltek

Az ortogonalis, szimplex és biortogonalis jeltek olyan példák, amelyek jól használhatók a teljesítménykorlátozott tartományban, amikor a sáv szélességre nincsen megkötés. Ezek között az ortogonalis jeltér tárgyalása és analízise a legegyszerűbb. A szimplex jeltér optimálisnak tűnik, ha a konstelláció mérete M adott, és nincsen megkötés a dimenziók számára. A biortogonalis jeltér kis mértékben sávtakarékosabb a másik kettőnél. Nagy M -ek esetén a három jeltér lényegében ekvivalens egymással.

A következő példákkal e három jeltér jellegzetes tulajdonságait elemezzük, és megmutatjuk, hogy ezek az uniós korlát szerint képesek az abszolút Shannon-korlátot 3 dB hibával megközelíteni. Az ilyen jeltek hátránya az, hogy a jeltér méretének a növelésével a dimenziók száma nő, és ha $M \rightarrow \infty$, akkor a $\rho(\mathcal{A})$ spektrális hatékonyság minden határon túl csökken.

- 1. példa: Ortogonalis jeltek

Ortogonalis jeltek esetén a jeltér M darab ortogonalis, \mathbb{R}^N -beli vektor halmaza, ahol $\mathcal{A} = \{\mathbf{a}_j, 1 \leq j \leq M\}$, és minden vektorpár esetén igaz, hogy $\langle \mathbf{a}_i, \mathbf{a}_j \rangle = 0$, ha $i \neq j$, és $\langle \mathbf{a}_i, \mathbf{a}_i \rangle = \|\mathbf{a}_i\|^2 = E(\mathcal{A})$ minden i -re, azaz példánkban legyenek a jelek azonos $E(\mathcal{A})$ energiájúak.

A korábbi ismeretek alapján a rendszer paraméterei az alábbi egyenletekkel adhatók meg, mivel $N = M$:

$$d_{min}^2(\mathcal{A}) = \|\mathbf{a}_i - \mathbf{a}_j\|^2 = \langle (\mathbf{a}_i - \mathbf{a}_j), (\mathbf{a}_i - \mathbf{a}_j) \rangle = \|\mathbf{a}_i\|^2 + \|\mathbf{a}_j\|^2 = 2E(\mathcal{A}),$$

$$\rho(\mathcal{A}) = \frac{2}{M} \log_2 M,$$

$$E_s(\mathcal{A}) = \frac{2}{M} E(\mathcal{A}),$$

$$E_b(\mathcal{A}) = \frac{E_s(\mathcal{A})}{\rho(\mathcal{A})} = \frac{E(\mathcal{A})}{\log_2 M},$$

$$K_{\min}(\mathcal{A}) = M - 1,$$

$$\begin{aligned} \Pr(e)(\mathcal{A}) &\leq (M-1)Q\left(\sqrt{\frac{E(\mathcal{A})}{N_0}}\right) = (M-1)Q\left(\sqrt{\frac{E_b \log_2 M}{N_0}}\right) \leq \\ &\leq (M-1) \exp\left(-\frac{E_b}{2N_0} \log_2 M\right) = (M-1) \exp\left(-\frac{E_b}{2N_0} \frac{\ln M}{\ln 2}\right) \leq \\ &\leq \exp(\ln M) \exp\left(-\frac{E_b}{2N_0} \frac{\ln M}{\ln 2}\right) = \exp\left(\ln M \left(1 - \frac{E_b}{2N_0 \ln 2}\right)\right). \end{aligned}$$

Ebből nyilvánvaló, hogy ha $\frac{E_b}{2N_0 \ln 2} > 1$, azaz $\frac{E_b}{N_0} > 2 \ln 2$ esetén M növelésével a hibaarány minden határon túl csökken. A jeltér tehát az uniós korlát alkalmazásával 3 dB-lel megközelíti az abszolút Shannon-korlátot. Megjegyzendő, hogy az ortogonális jeltérrel a valóságban (uniós korlát nélkül) a Shannon-korlát el is érhető, természetesen a nominális spektrális hatékonyság ekkor is nullához tart.

• 2. példa: Szimplex jelterek

Az \mathcal{A}' szimplex jeltér az 1. példában szereplő \mathcal{A} ortogonális jeltérből származtatható az ortogonális jeltér szimplexének előállításával. Ehhez definiáljuk az \mathcal{A} jeltér $\mathbf{m}(\mathcal{A})$ átlagát, ami az

$$\mathbf{m}(\mathcal{A}) = \frac{\sum_{j=1}^M \mathbf{a}_j}{M}$$

egyenlettel határozható meg, amiről biztosan tudjuk, hogy nem lehet nulla, mivel ellenkező esetben fennállna az

$$\mathbf{a}_i = - \sum_{\substack{j=1 \\ j \neq i}}^M \mathbf{a}_j$$

összefüggés, ami ellentmondásban van az $\{\mathbf{a}_j\}$ jelek ortogonalitásával.

A fentiek alapján az $\mathcal{A}' = \{\mathbf{a}'_j, 1 \leq j \leq M\}$ szimplex jeltér j -dik elemét az $\mathbf{a}'_j = \mathbf{a}_j - \mathbf{m}(\mathcal{A})$ kifejezéssel adhatjuk meg. A definícióból nyilvánvaló, hogy a jelek közötti euklideszi távolságok az eredeti ortogonális jeltérben és a szimplex jeltérben azonosak, mivel $(\mathbf{a}'_j - \mathbf{a}'_i) = (\mathbf{a}_j - \mathbf{a}_i)$.

A korábbi ismeretek alapján a szimplex jeltér paraméterei az alábbi egyenletekkel adhatók meg:

$$d_{\min}^2(\mathcal{A}') = \|\mathbf{a}'_i - \mathbf{a}'_j\|^2 = \|\mathbf{a}_i - \mathbf{a}_j\|^2 = d_{\min}^2(\mathcal{A}) = 2E(\mathcal{A}),$$

$$\rho(\mathcal{A}') = \rho(\mathcal{A}) = \frac{2}{M} \log_2 M,$$

$$E_s(\mathcal{A}') = \frac{2}{M} E(\mathcal{A}'),$$

ahol

$$E(\mathcal{A}') = \frac{1}{M} \sum_{j=1}^M \|\mathbf{a}'_j\|^2 = \frac{1}{M} \sum_{j=1}^M \|\mathbf{a}_j - \mathbf{m}(\mathcal{A})\|^2 = E(\mathcal{A}) - \frac{E(\mathcal{A})}{M},$$

mivel

$$\|\mathbf{a}_j - \mathbf{m}(\mathcal{A})\|^2 = \langle (\mathbf{a}_j - \mathbf{m}(\mathcal{A})), (\mathbf{a}_j - \mathbf{m}(\mathcal{A})) \rangle = \|\mathbf{a}_j\|^2 - 2\langle \mathbf{a}_j, \mathbf{m}(\mathcal{A}) \rangle + \|\mathbf{m}(\mathcal{A})\|^2,$$

$$\|\mathbf{a}_j\|^2 = E(\mathcal{A}),$$

$$\langle \mathbf{a}_j, \mathbf{m}(\mathcal{A}) \rangle = \langle \mathbf{a}_j, \frac{\sum_{j=1}^M \mathbf{a}_j}{M} \rangle = \frac{E(\mathcal{A})}{M}$$

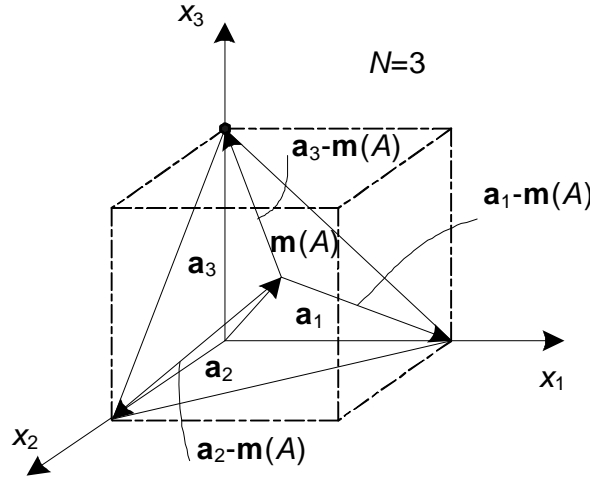
és

$$\|\mathbf{m}(\mathcal{A})\|^2 = \frac{\langle (\sum_{j=1}^M \mathbf{a}_j), (\sum_{k=1}^M \mathbf{a}_k) \rangle}{M^2} = \frac{M E(\mathcal{A})}{M^2} = \frac{E(\mathcal{A})}{M}.$$

Ezen kívül tudjuk, hogy ha $i \neq j$, akkor

$$\begin{aligned} \langle \mathbf{a}'_i, \mathbf{a}'_j \rangle &= \langle (\mathbf{a}_i - \mathbf{m}(\mathcal{A})), (\mathbf{a}_j - \mathbf{m}(\mathcal{A})) \rangle = \\ &= \langle \mathbf{a}_i, \mathbf{a}_j \rangle - \langle \mathbf{a}_i, \mathbf{m}(\mathcal{A}) \rangle - \langle \mathbf{a}_j, \mathbf{m}(\mathcal{A}) \rangle + \|\mathbf{m}(\mathcal{A})\|^2 = 0 - 2\frac{E(\mathcal{A})}{M} + \frac{E(\mathcal{A})}{M} = -\frac{E(\mathcal{A})}{M}. \end{aligned}$$

Az eredeti ortogonális és a belőle származtatott szimplex jelretet a 6.2. ábrán mutatjuk be $N = 3$ esetén.



6.2. ábra. A háromdimenziós szimplex jelret illusztrációja

- 3. példa: Biortogonális jeltek

A biortogonális jelteket az 1. példában szereplő $\mathcal{A} = \{\mathbf{a}_j, 1 \leq j \leq M\}$ ortogonális jeltérből származtathatjuk az $\mathcal{A}'' = \pm \mathcal{A}$ szabály szerint, azaz a biortogonális jeltér tartalmazza az eredeti ortogonális jeltér minden elemét és azok mínusz egyszeresét is. A biortogonális jeltér elemeinek a száma $M'' = 2M$ és átlagos szimbólum energiája $E(\mathcal{A}'') = E(\mathcal{A})$.

A korábbi ismeretek alapján a rendszer paraméterei az alábbi egyenletekkel adhatók meg:

$$d_{min}^2(\mathcal{A}'') = d_{min}^2(\mathcal{A}) = 2E(\mathcal{A}),$$

$$\mathbf{m}(\mathcal{A}'') = 0, \quad N'' = N = M, \quad M'' = 2M,$$

$$\rho(\mathcal{A}'') = \frac{2}{N} \log_2 M'' = \frac{2}{M} \log_2 2M = \frac{2}{M} (1 + \log_2 M),$$

és ha $M'' = 2M = 2^k$, akkor

$$\rho(\mathcal{A}'') = \frac{2}{2^{k-1}} \log_2 2^k = \frac{2}{2^{k-1}} k = 4k2^{-k}$$

bit/két dimenzió.

Ezt felhasználva a biortogonális jeltér névleges kódolási nyeresége a

$$\gamma_c(\mathcal{A}'') = \frac{d_{\min}^2(\mathcal{A}'')}{4E_b(\mathcal{A}'')} = \frac{2E(\mathcal{A})}{4\frac{E(\mathcal{A})}{k}} = \frac{2kE(\mathcal{A})}{4E(\mathcal{A})} = \frac{k}{2},$$

mivel

$$E_b(\mathcal{A}'') = \frac{E_s(\mathcal{A}'')}{\rho(\mathcal{A}'')} = \frac{E(\mathcal{A})}{\log_2 2M} = \frac{E(\mathcal{A})}{k},$$

a legközelebbi szomszédok átlagos száma

$$K_{\min}(\mathcal{A}'') = 2M - 2 = 2^k - 2.$$

Ezeket az eredményeket felhasználva vizsgáljunk meg néhány számpéldát különböző biortogonális jelkészletek esetén (a számpéldáknál a '' jelöléseket elhagyjuk, azaz a paraméterek a vizsgált biortogonális jelkészletre vonatkoznak)

a.) Legyen $k = 4$, $N = 2^{k-1} = 8$, tehát a jeltér mérete $M = 2^k = 16$, ez azt jelenti, hogy szimbólumonként $\log_2 M = 4$ bit információt viszünk át a csatornán. Tételezzük fel, hogy $E(\mathcal{A})$ adott.

A rendszer paraméterei a következők:

$$\begin{aligned} \rho &= \frac{2}{N} \log_2 M = \frac{2}{2^{k-1}} \log_2 2^k = 1 \text{ [bit/két dimenzió]}, & E_b &= \frac{E(\mathcal{A})}{\log_2 M} = \frac{E(\mathcal{A})}{4}, \\ K_b(\mathcal{A}) &= \frac{K_{\min}}{\log_2 M} = \frac{2^k - 2}{k} = 3,5, & d_{\min}^2(\mathcal{A}) &= 2E(\mathcal{A}), & \gamma_c(\mathcal{A}) &= \frac{k}{2} = 2 \text{ [3,01 dB]}, \\ \gamma_{eff}(\mathcal{A}) &\approx 3 - 0,2(\log_2 K_b(\mathcal{A})) = 2,6 \text{ [dB]}. \end{aligned}$$

b.) Legyen $k = 6$, $N = 2^{k-1} = 32$, tehát a jeltér mérete $M = 2^k = 64$, ez azt jelenti, hogy szimbólumonként $\log_2 M = 6$ bit információt viszünk át a csatornán. Tételezzük fel, hogy $E(\mathcal{A})$ adott.

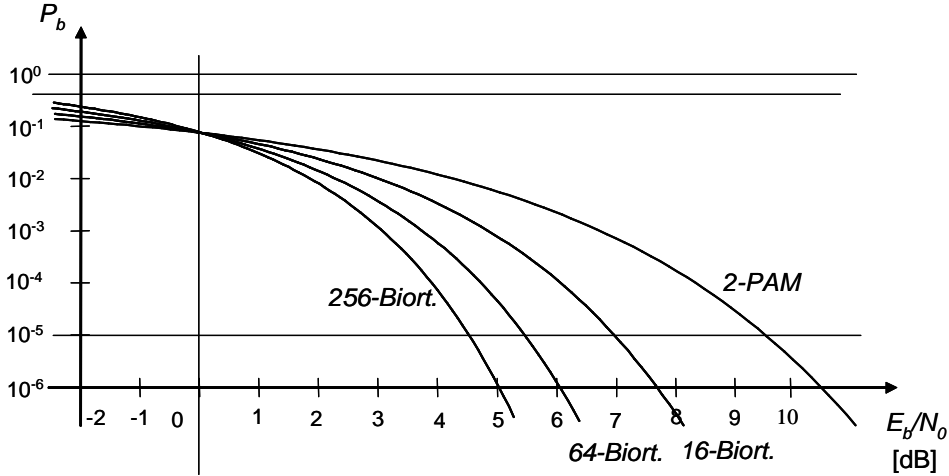
A rendszer paraméterei a következők:

$$\begin{aligned} \rho &= \frac{2}{2^{k-1}} \log_2 2^k = \frac{12}{32} \text{ [bit/két dimenzió]}, & E_b &= \frac{E(\mathcal{A})}{6}, \\ K_b(\mathcal{A}) &= \frac{2^k - 2}{k} = \frac{62}{6}, & \gamma_c(\mathcal{A}) &= \frac{k}{2} = 3 \text{ [4,8 dB]}, \\ \gamma_{eff}(\mathcal{A}) &\approx 4,8 - 0,2(\log_2 K_b(\mathcal{A})) = 4,1 \text{ [dB]}. \end{aligned}$$

c.) Legyen $k = 8$, $N = 2^{k-1} = 128$, tehát a jeltér mérete $M = 2^k = 256$, ez azt jelenti, hogy szimbólumonként $\log_2 M = 8$ bit információt viszünk át a csatornán. Tételezzük fel, hogy $E(\mathcal{A})$ ismét adott.

A rendszer paraméterei a következők:

$$\begin{aligned} \rho &= \frac{2}{2^{k-1}} \log_2 2^k = \frac{28}{128} \text{ [bit/két dimenzió]}, & E_b &= \frac{E(\mathcal{A})}{8}, \\ K_b(\mathcal{A}) &= \frac{2^k - 2}{k} = \frac{254}{8}, & \gamma_c(\mathcal{A}) &= \frac{k}{2} = 4 \text{ [6 dB]}, \\ \gamma_{eff}(\mathcal{A}) &\approx 6 - 0,2(\log_2 K_b(\mathcal{A})) = 5 \text{ [dB]}. \end{aligned}$$



6.3. ábra. A P_b bit hibaarány az E_b/N_0 függvényében a $2^k = 16, 64$ és 256 paraméterű biortogonális jelkészletek esetén

A 6.3. ábrán megadtuk a fentebb vizsgált biortogonális jelkészletek bithibaarányának uniós korlátját a jel-zaj viszony függvényében. Megállapítható, hogy az alkalmazott közelítés $P_b = 10^{-5}$ értéknél meglehetősen pontos.

6.3. A teljesítőképesség vizsgálata a sávkorlátozott tartományban

A sávkorlátozott tartomány korábbi definíciója szerint, sávkorlátozott rendszerről akkor beszélünk, ha a névleges spektrális hatékonyság $\rho \geq 2$ [bit/két dimenzió], azaz ebben a tartományban nem bináris modulációs rendszert alkalmazunk. Itt minden jellemző mennyiséget két dimenzióra normálunk, és az SNR_{norm} értéket tekintjük a jel-zaj viszony normalizált értékének. A teljesítőképesség szempontjából a kódolatlan M -PAM jelkészletet $\mathcal{A} = \alpha\{\pm 1, \pm 3, \dots, \pm(M-1)\}$, vagy a vele ekvivalens $(M \times M)$ -QAM (\mathcal{A}^2) konstellációt tekintjük összehasonlítási alapnak. Referenciarendszereinkben M tipikusan 2 hatványa. Ezért az ilyen konstellációk névleges spektrális hatékonysága $\rho = 2 \log_2 M$ [bit/két dimenzió], a minimális négyzetes távolság pedig $d_{min}^2(\mathcal{A}^2) = 4\alpha^2$. A korábbi fejezetekben megmutattuk, hogy a két dimenzióra jutó átlagos energia ilyen rendszerekben

$$E_s = \frac{2\alpha^2(M^2 - 1)}{3} = \frac{d_{min}^2(\mathcal{A})(2^p - 1)}{6}.$$

A legközelebbi szomszédok átlagos száma két dimenzióként $K_s(\mathcal{A})$ kétszerese az M -PAM rendszerben mért értéknek, azaz $K_s(\mathcal{A}) = 4(M-1)/M$, ami M növekedésével gyorsan a $K_s(\mathcal{A}) \approx 4$ értékhez tart. Az uniós korlát szerint a két dimenzióra eső hibaarány a

$$P_s(e) \approx 4Q\left(\sqrt{3SNR_{norm}}\right)$$

kifejezéssel közelíthető. A $P_s(e)$ alapgörbéje az SNR_{norm} függvényében az 5.5. ábrán látható.

A sávkorlátozott rendszerekben az \mathcal{A} jeltér γ_{eff} effektív kódolási nyeresége definíciószerűen az adott \mathcal{A} és az M -PAM (vagy a kódolatlan $M \times M$ -QAM) jeltérben egy adott $P_s(e)$ hibaarány eléréséhez szükséges SNR_{norm} értékek dB-ben kifejezett értékeinek a különbsége.

Például az 5.5. ábra szerint a $P_s(e) \approx 10^{-5}$ értékhez maximálisan 8,4 dB effektív kódolási nyereség tartozik, ami kb. 3 dB értékkel kisebb, mint a teljesítménykorlátozott tartományban (annak következtében, hogy itt a rendszer sávszélessége korlátozott).

Az effektív kódolási nyereséget az uniós korlát alkalmazásával a következőképpen határozhatjuk meg. Tudjuk, hogy a két dimenzióra eső hibaarány a

$$P_s(e) = \frac{2\Pr(e)}{N} \approx \frac{2K_{\min}(\mathcal{A})}{N} Q\left(\sqrt{\frac{d_{\min}^2(\mathcal{A})}{2N_0}}\right)$$

kifejezéssel határozható meg, és

$$SNR_{norm} = \frac{\frac{E_s}{N_0}}{2^p - 1},$$

amiből

$$\frac{d_{\min}^2(\mathcal{A})}{2N_0} = \frac{d_{\min}^2(\mathcal{A})(2^p - 1)SNR_{norm}}{2E_s}.$$

Ezért sávkorlátozott tartományban a $\gamma_c(\mathcal{A})$ nominális kódolási nyereség:

$$\gamma_c(\mathcal{A}) = \frac{(2^p - 1)d_{\min}^2(\mathcal{A})}{6E_s},$$

és ez a definíció úgy van normálva, hogy M -PAM vagy $(M \times M)$ -QAM esetén, $\gamma_c(\mathcal{A}) = 1$. Egyébként $\gamma_c(\mathcal{A})$ -t általában dB-ben mérjük. Ezután az uniós korlát alkalmazásával

$$P_s(e) \approx K_s(\mathcal{A})Q\left(\sqrt{3\gamma_c(\mathcal{A})SNR_{norm}}\right),$$

ahol $K_s(\mathcal{A}) = 2K_{\min}(\mathcal{A})/N$ a két dimenzióra eső legközelebbi szomszédok átlagos száma. Jegyezzük meg, hogy M -PAM vagy $(M \times M)$ -QAM esetén ez a kifejezés a $P_s(e) \approx 4Q(\sqrt{3SNR_{norm}})$ formára egyszerűsödik.

Adott $\gamma_c(\mathcal{A})$ és $K_s(\mathcal{A})$ esetén a hibaarány az 5.5. ábra alapján az alábbi módon határozható meg az uniós korlát alkalmazásával: az ábrán megadott alapgörbét dB-ben $\gamma_c(\mathcal{A})$ értékkel balra, és $K_s(\mathcal{A})/4$ értékkel felfelé kell eltolni. A korábban megismert ökölszabály most is érvényes, miszerint egy kettes faktor a $K_s(\mathcal{A})$ értékében kb. 0,2 dB változást okoz az effektív kódolási nyereségben $P_s(e) \approx 10^{-5}$ értéknél, ha a $K_s(\mathcal{A})$ nem túlságosan nagy értékű.

6.4. Összefoglalás

A fejezet eredményei az alábbiakban adhatók meg:

- A teljesítménykorlátozott tartományban a névleges kódolási nyereség $\gamma_c(\mathcal{A}) = d_{\min}^2(\mathcal{A})/4E_b$, és az uniós korlát pontosságával $P_b(e) \approx K_b(\mathcal{A})Q\left(\sqrt{2\gamma_c(\mathcal{A})E_b/N_0}\right)$. Ezt a görbét meg lehet szerkeszteni úgy, hogy az 5.4. ábrán megadott $P_b(e) \approx Q\left(\sqrt{2E_b/N_0}\right)$ referenciadiagramot balra toljuk el $\gamma_c(\mathcal{A})$ értékkel és felfelé toljuk el $K_b(\mathcal{A})$ értékkel dB-ben. $P_b(e) \approx 10^{-5}$ értéknél az effektív kódolási nyereség a $\gamma_{eff}(\mathcal{A}) \approx \gamma_c(\mathcal{A}) - (0,2)(\log_2 K_b(\mathcal{A}))$ kifejezéssel közelíthető.
- A sávkorlátozott tartományban a névleges kódolási nyereség $\gamma_c(\mathcal{A}) = (2^p - 1)d_{\min}^2(\mathcal{A})/6E_s$, és az uniós korlát pontosságával $P_s(e) \approx K_s(\mathcal{A})Q\left(\sqrt{3\gamma_c(\mathcal{A})SNR_{norm}}\right)$. Ezt a görbét meg lehet szerkeszteni úgy, hogy az 5.5. ábrán megadott $P_s(e) \approx 4Q(\sqrt{3SNR_{norm}})$ referenciadiagramot balra toljuk el $\gamma_c(\mathcal{A})$ értékkel és felfelé toljuk el $K_s(\mathcal{A})/4$ értékkel dB-ben. $P_s(e) \approx 10^{-5}$ értéknél az effektív kódolási nyereség a $\gamma_{eff}(\mathcal{A}) \approx \gamma_c(\mathcal{A}) - (0,2)\log_2(K_s(\mathcal{A})/4)$ kifejezéssel közelíthető.

7. fejezet

Bevezetés a bináris kódok világába

Ebben a fejezetben azokat a bináris jelkonstellációkat vizsgáljuk, melyekhez bináris blokk kódok tartoznak. Az ilyen konstellációk esetén a bitsebesség (a névleges spektrális hatékonyság) $\rho \leq 2$ [bit/két dimenzió], ezért ezek tipikusan a teljesítménykorlátozott tartományban működnek. A fejezetben a bináris lineáris blokk kódokkal fogunk foglalkozni, amelyek meglehetősen érdekes és sajátos algebrai struktúrával rendelkeznek. Ugyanis ezek vektor teret képeznek az \mathbb{F}_2 bináris tér felett. Az ilyen kódok egy jól használható végtelen elemű családja a Reed-Muller kódok halmaza. Tárgyalni fogjuk, hogy az úgynevezett kemény döntés (hard decision) alkalmazása mekkora hibát okoz, elemezzük a hibajavítás hatását, megmutatjuk, hogy milyen hatása van törlés (erasure) alkalmazásának, és foglalkozunk az általánosított minimum távolságú dekódolással.

7.1. Bináris jelkonstellációk

Ebben a fejezetben azokat a konstellációkat vizsgáljuk, amelyeknél a bináris kódok euklideszi térbeli vetülete koordinátáinként a 2-PAM elrendezésnek felel meg. Az ilyen konstellációkat bináris jelkonstellációknak nevezzük. Egy n hosszúságú bináris blokk kód $C \subseteq \{0, 1\}^n$ az n hosszúságú bináris sorozatok egy tetszőleges részhalmaza. A továbbiakban a $\{0, 1\}$ bináris ABC-t azonosnak fogjuk tekinteni a két elemű \mathbb{F}_2 térrel, amire érvényesek a mod-2 aritmetikai szabályok. Ezen kívül leszögezzük azt a követelményt, hogy a C kód legyen lineáris, azaz C egy altere az összes n hosszúságú bináris vektorok \mathbb{F}_2^n terének.

Az alábbiakban röviden összefoglaljuk a fent említett algebrai tulajdonságokat. Egy $\mathbf{x} \in C$ bináris kódszó minden $x_k \in \{0, 1\}$ elemét leképezzük az $\mathcal{A} = \{\pm\alpha\}$ 2-PAM jeltér egy pontjára az $\{s : \{0, 1\} \rightarrow \mathcal{A}\}$ függvény segítségével. Ennek két szokásos kifejezése van:

$$s(x) = \alpha(-1)^x$$

és

$$s(x) = \alpha(1 - 2x).$$

Az első azt fejezi ki, hogy izomorfizmus áll fent az additív $\mathbb{Z}_2 = \{0, 1\}$ bináris csoport és a multiplikatív $\{\pm 1\}$ csoport között, mivel $s(x)s(x') = (-1)^{x+x'}$. A második pedig egy egyszerű transzformáció, amely a bináris $\{0, 1\} \in \mathbb{F}_2$ tér és a valós $\{-1, 1\} \in \mathbb{R}$ tér között teremt kapcsolatot. Mindkét reprezentáció ugyanahhoz az

$$s(0) = \alpha \quad s(1) = -\alpha$$

leképezéshez vezet.

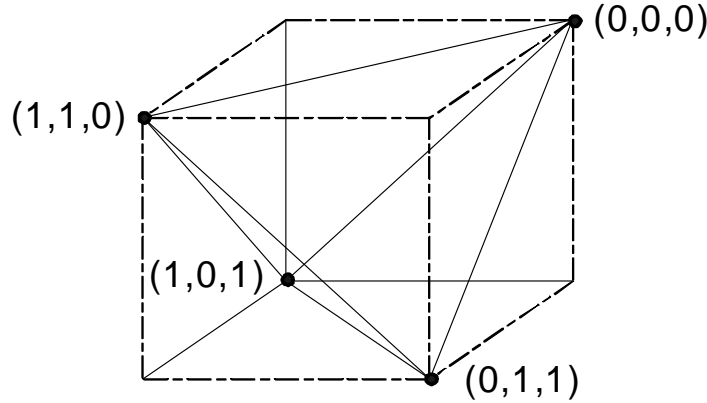
A 2-PAM leképezés szerint az $(\mathbb{F}_2)^n$ -beli n hosszúságú sorozatokat a valós $\{\pm\alpha, \pm\alpha, \dots, \pm\alpha\}$ n hosszúságú sorozatokba képezzük le. Geometriai értelemben ez nem jelent mást, mint azt, hogy az egyes kódszavakat az origó középpontú, 2α élhosszúságú n -dimenziós kocka csúcsaihoz rendeljük.

Ebből az következik, hogy az $\mathcal{A}' = s(C)$ konstelláció nem más, mint a $C \subseteq (\mathbb{F}_2)^n$ kód leképezése az n -dimenziós kocka csúcsaira.

Egy n -dimenziós bináris \mathcal{A} konstelláció mérete az $|\mathcal{A}| \leq 2^n$ értékkel felülről korlátozott, és a hozzá tartozó $\rho = 2/n \log_2 2^n$ bitsebesség felső korlátja $\rho \leq 2$ [bit/két dimenzió]. Ezek szerint az ilyen bináris konstellációkat csak a teljesítménykorlátozott tartományban lehet használni.

Mivel az n -dimenziós kocka típusú $\mathcal{A}^n = s((\mathbb{F}_2)^n) = (s(\mathbb{F}_2))^n$ konstelláció a 2-PAM konstelláció n -szeres Descartes szorzata, ezért a normalizált paraméterei azonosak a 2-PAM rendszerével, és nem rendelkezik kódolási nyereséggel. Ennek ellenére abban bízhatunk, hogy az \mathcal{A}^n egy \mathcal{A}' részhalmaza a használata esetén az euklideszi távolság növelésével mégis el tudunk érni egy adott kódolási nyereséget.

Példa



7.1. ábra. A $C = \{000, 011, 110, 101\}$ kód euklideszi leképezése az \mathbb{R}^3 térbe

A 7.1. ábrán megadtuk a $C = \{000, 011, 110, 101\}$ kód euklideszi leképezését a valós számok háromdimenziós térébe. A C kód négy három hosszúságú bináris kódszóból áll. A kód bitsebessége $\rho = 4/3$ [bit/két dimenzió]. A kódszavakat az $s(C)$ transzformációval a három dimenziós 2α élhosszúságú kocka négy csúcspontjához rendeltük. A kód minimális négyzetes euklideszi távolsága $d_{min}^2(s(C)) = 8\alpha^2$, az $s(C)$ leképezés minden jelvektorának 3 minimális távolságra eső szomszédja van. A jeltér átlagos energiája $E(s(C)) = 3\alpha^2$, ezért az átlagos bitenergia $E_b = (3/2)\alpha^2$, és a rendszer nominális kódolási nyeresége a

$$\gamma_c(s(C)) = \frac{d_{min}^2(s(C))}{4E_b(s(C))} = \frac{4}{3} \quad (1.25) \text{ [dB]},$$

kifejezéssel adható meg.

Az eddigiekből nyilvánvaló, hogy az a megkötés, hogy a C kód minden elemét az n -dimenziós kocka csúcsaihoz rendeltük, biztosan erősen szuboptimális. Ugyanakkor a vizsgálatokból kiderül, hogy ez a komoly megkötés a rendszer teljesítőképességét jelentősen nem befolyásolja. A véletlen kódolások kapacitászámítások szerint az $\mathcal{A} = \{\pm\alpha\}$ típusú ABC alkalmazásakor a Shannon-korlát $\rho = 1$ [bit/két dimenzió] értéknél 0 dB helyett ilyenkor 0,2 dB, azaz a veszteség 0,2 dB, de ρ csökkenésével a veszteség nullához tart. Ez azt jelenti, hogy a bináris jelkonstellációk (a bináris blokk kódok) kis ρ esetén elegendően hatékonyak.

7.2. A bináris lineáris blokk kódok, mint bináris vektorterek

Az általunk vizsgált bináris kódok mindegyike lineáris kód. Egy bináris lineáris blokk kód az n hosszú bináris sorozatoknak egy olyan halmaza, amelyeknek minden összetevője eleme az $\mathbb{F}_2 = \{0, 1\}$ bináris véges testnek, azaz a kód az \mathbb{F}_2 bináris véges test feletti vektorok halmaza.

Amint azt a következőkben látni fogjuk a C kód csoportot kell képezzen a n hosszúságú bináris sorozatok összeadására nézve. Éppen ezért először az $\mathbb{F}_2 = \{0, 1\}$ bináris test és az \mathbb{F}_2 test feletti vektortér algebrai tulajdonságait vizsgáljuk meg.

Általában egy \mathbb{F} test olyan elemek halmaza, melyekre értelmezhető az összeadás és szorzás művelete, és amelyek érvényesek a szokásos aritmetikai szabályok (kommutativitás, asszociativitás, disztributivitás). A testnek van 0 nulleleme (additív neutrális eleme), melyre igaz, hogy $a + 0 = a$ minden $a \in \mathbb{F}$, és minden a testelemhez tartozik egy additív inverz $-a$, úgy, hogy $(-a) + a = 0$. A testnek van 1 egységeleme (multiplikatív neutrális eleme), melyre igaz, hogy $1 * a = a$ minden $a \in \mathbb{F}$, és minden a testelemhez tartozik egy multiplikatív inverz a^{-1} , úgy, hogy $(a^{-1}) * a = 1$.

Az \mathbb{F}_2 bináris test (ezt sokszor Galois testnek is szoktuk nevezni, és $GF(2)$ -nek jelöljük) csak két elemet (0-t és 1-et) tartalmaz, ami például a páros és páratlan egész számok moduló 2-es maradékát jelentheti. Az összeadásra és szorzásra vonatkozó táblázat a mod 2 aritmetikára épül, ahol 0 az additív, 1 a multiplikatív egységelem:

$$\begin{array}{ll} 0 + 0 = 0 & 0 * 0 = 0, \\ 0 + 1 = 1 & 0 * 1 = 0, \\ 1 + 0 = 1 & 1 * 0 = 0, \\ 1 + 1 = 0 & 1 * 1 = 1. \end{array}$$

Valójában ezek a szabályok az 0 és 1 elemekre minden test esetén érvényesek. Jegyezzük meg, hogy az 1 additív inverze 1, így minden $a \in \mathbb{F}$ elem inverze maga az a ($-a = a$).

Általában egy \mathbb{F} feletti \mathcal{V} vektortér olyan \mathbf{v} vektorok halmaza, melyek között szerepel a $\mathbf{0}$ elem, és amelyekre értelmezhető a vektorok összeadása és skalárral való szorzása, és teljesülnek más vektortérre vonatkozó axiómák.

Az \mathbb{F}_2 test felett értelmezett \mathcal{V} vektortér esetén a skalárral való szorzás triviálisan értelmezhető, mivel $0 * \mathbf{v} = \mathbf{0}$ és $1 * \mathbf{v} = \mathbf{v}$ automatikusan eleme \mathcal{V} -nek. Így mindössze az additivitási feltétel (csoport tulajdonság) kell, hogy teljesüljön: ha $\mathbf{v}, \mathbf{v}' \in \mathcal{V}$, akkor $\mathbf{v} + \mathbf{v}' \in \mathcal{V}$. Végezetül megállapíthatjuk, hogy minden $\mathbf{v} \in \mathcal{V}$ elemnek van additív inverze, mivel

$$\mathbf{v} + \mathbf{v} = 1 * \mathbf{v} + 1 * \mathbf{v} = (1 + 1) * \mathbf{v} = 0 * \mathbf{v} = \mathbf{0},$$

ezért a bináris test felett az összeadás és a kivonás azonos művelet.

Az \mathbb{F}_2 test feletti vektorteret bináris vektortérnek nevezzük. Elemi példa erre az $(\mathbb{F}_2)^n$ halmaz, ami az összes $\mathbf{v} = (v_1, \dots, v_n)$ n hosszúságú bináris sorozatok halmaza, komponensenkénti mod-2 összeadással és a skalárral való szorzással. A továbbiakban minden esetben olyan bináris vektorterekről fogunk beszélni, melyek $C \subseteq (\mathbb{F}_2)^n$ alterei az $(\mathbb{F}_2)^n$ vektortérnek, és amiket n hosszúságú C bináris blokk kódnak fogunk nevezni.

Ha $\mathcal{G} = (\mathbf{g}_1, \dots, \mathbf{g}_k)$ bináris vektorok halmaza a \mathcal{V} vektortérben, akkor $C(\mathcal{G})$ ezen vektorok összes lineáris kombinációja:

$$C(\mathcal{G}) = \left\{ \sum_j a_j \mathbf{g}_j, a_j \in \mathbb{F}_2, 1 \leq j \leq k \right\}$$

ami a \mathcal{V} altere, mivel $C(\mathcal{G})$ nyilvánvalóan rendelkezik a csoport tulajdonsággal. A \mathcal{G} halmaz lineárisan független, ha mind a 2^k lineáris kombináció különbözik egymástól, azaz a $C(\mathcal{G})$ mérete $|C(\mathcal{G})| = 2^k$. Egy olyan lineárisan független vektorokból álló \mathcal{G} halmazzal, melyekre igaz, hogy $C(\mathcal{G}) = \mathcal{V}$ a \mathcal{V} vektortér bázisának nevezzük, és a bázis $\{\mathbf{g}_1, \dots, \mathbf{g}_k\}$ elemei a generátorok. A $\mathcal{G} = (\mathbf{g}_1, \dots, \mathbf{g}_k)$ halmazzal egy $k \times n$ méretű \mathbf{G} mátrixba is rendezhetjük az \mathbb{F}_2 felett, amit a $C(\mathcal{G})$ generátormátrixának nevezünk.

A \mathcal{V} vektortér dimenziója k , a generátorok száma, annak bármely bázisát tekintve. Bármely \mathcal{V} vektortér esetén a vektortér \mathcal{G} bázisát és k dimenzióját az alábbi mohó algoritmussal határozhatjuk meg:

- Kezdeti feltétel: legyen $k = 0$ és $\mathcal{G} = O$ (üres halmaz);

- A ciklus kezdete: ha $C(\mathcal{G}) = \mathcal{V}$, akkor a ciklusnak vége, és $\dim(\mathcal{V}) = k$; egyébként növeljük meg k értékét eggyel, és tekintjük \mathbf{g}_k -nak az $\mathbf{v} \in (\mathcal{V} - C(\mathcal{G}))$ halmaz egyik elemét;
- Térjünk vissza a ciklus elejére.

Ily módon a \mathcal{V} vektortér mérete mindig $|\mathcal{V}| = 2^k$ és $\dim(\mathcal{V}) = k$, vagy másképpen $\dim(\mathcal{V}) = \log_2 |\mathcal{V}|$.

Egy (n, k) bináris lineáris C kód az $(\mathbb{F}_2)^n$ vektortér egy k dimenziós altere, melynek a mérete 2^k . Más szóval egy (n, k) bináris lineáris kód az n hosszúságú bináris vektorok 2^k elemű halmaza, amely tartalmazza a $\mathbf{0}$ elemet, és a konvencionális bináris összeadásra nézve csoportot alkot.

Példák (egyszerű bináris lineáris kódok)

Az (n, n) bináris lineáris kód az $(\mathbb{F}_2)^n$ vektortér összes n hosszúságú bináris vektorát tartalmazza. Ezt a kódot gyakran n hosszúságú univerzális kódnak nevezzük.

Az $(n, 0)$ bináris lineáris kód a $C = \{O\}$, amely egyetlen csupa nulla elemet tartalmazó n hosszúságú bináris vektorból áll. Ezt gyakran n hosszúságú triviális kódnak nevezzük.

Az $(n, 1)$ bináris lineáris kódot, amely az n hosszúságú csupa 0 és csupa 1 elemű bináris vektorokból álló n hosszúságú ismétléses kódnak nevezzük.

Az $(n, n-1)$ bináris lineáris kód az összes n hosszúságú bináris vektorokból áll, melyekben páros számú 1-es található. Az ilyen kódokat páros súlyú, vagy paritás ellenőrző (single-parity-check, SPC) kódnak nevezzük.

A Hamming metrika

Az $(\mathbb{F}_2)^n$ bináris vektortér geometriája a Hamming metrikával jellemezhető:

$$w_H(\mathbf{x}) = \text{az egyesek száma } (\mathbf{x}) \text{ -ben,}$$

ahol $w_H(\mathbf{x})$ az \mathbf{x} vektor Hamming súlya.

A Hamming metrika teljesíti a metrikára érvényes általános axiómákat:

- Szigorúan pozitív: $w_H(\mathbf{x}) \geq 0$, és akkor és csak akkor áll fent az egyenlőség, ha $\mathbf{x} = \mathbf{0}$;
- Szimmetrikus: $w_H(-\mathbf{x}) = w_H(\mathbf{x})$, mivel $(-\mathbf{x}) = \mathbf{x}$;
- Érvényes rá a háromszög egyenlőtlenség: $w_H(\mathbf{x} + \mathbf{y}) \leq w_H(\mathbf{x}) + w_H(\mathbf{y})$.

Emiatt a

$$d(\mathbf{x}, \mathbf{y}) = w_H(\mathbf{x} - \mathbf{y}) = w_H(\mathbf{x} + \mathbf{y})$$

Hamming távolságot felhasználásával az $(\mathbb{F}_2)^n$ vektorteret metrikus térként lehet kezelni, ezt a teret Hamming térnek is szoktuk nevezni.

A továbbiakban megmutatjuk, hogy a C bináris lineáris blokk kódok csoport tulajdonsága következtében az egyes kódszavak közötti Hamming távolságok bármely kódszótól számítva azonosak. Ez azt jelenti, hogy bármely $\mathbf{x} \in C$ kódszót kiválasztva a többi kódszó \mathbf{x} -től mért távolságának eloszlása független \mathbf{x} megválasztásától.

Legyen $\mathbf{x} \in C$ egy adott C -beli kódszó, és vizsgáljuk meg az $\{(\mathbf{x} + \mathbf{y}) \mid \mathbf{y} \in C\} = \mathbf{x} + C$ halmazt, ha \mathbf{y} felveszi az összes lehetséges C -beli értéket. A C csoport-tulajdonsága miatt $\mathbf{x} + \mathbf{y}$ szintén kódszó C -ben. Sőt tudjuk, hogy a $(\mathbf{x} + \mathbf{y}) = (\mathbf{x} + \mathbf{y}')$ egyenlőség akkor és csak akkor áll fent, ha $\mathbf{y} = \mathbf{y}'$, mivel minden kódszó különbözik egymástól. Viszont az $\mathbf{x} + C$ halmaz mérete $|C|$, amiből következik, hogy $\mathbf{x} + C = C$, azaz $(\mathbf{x} + \mathbf{y})$ felveszi C minden elemét, ha \mathbf{y} felveszi C minden elemét. Mivel $d_H(\mathbf{x}, \mathbf{y}) = w_H(\mathbf{x} + \mathbf{y})$, a fentiekből az alábbi tétel következik:

Távolság invariancia tétel

Egy $\mathbf{x} \in C$ kódszótól mért $d_H(\mathbf{x}, \mathbf{y})$ Hamming távolságok halmaza, ha \mathbf{y} végigfut C minden elemén, független \mathbf{x} -től, és azonos a $d_H(\mathbf{0}, \mathbf{y})$ halmazzal, azaz a $\mathbf{0} \in C$ kódszótól mért Hamming távolságokkal, vagyis az összes $\mathbf{y} \in C$ kódszó Hamming súlyával.

Ha az (n, k) bináris lineáris C blokk kód minimális Hamming távolsága d , akkor a kódot (n, k, d) -vel jelöljük, azaz

$$d = \min_{\mathbf{x} \neq \mathbf{y} \in C} d_H(\mathbf{x}, \mathbf{y}).$$

A tétel alapján kimondhatunk egy segédtételt is:

A minimális Hamming távolság és a minimális Hamming súly egyenlőségének a segédtétele

A C kód minimális Hamming távolsága azonos a C -beli nem nulla vektorok minimális Hamming súlyával. Általánosabban: N_d , azon $\mathbf{y} \in C$ kódszavak száma amelyek egy $\mathbf{x} \in C$ kódszótól d távolságra vannak, független az \mathbf{x} megválasztásától.

Példák (Hamming távolságok)

Az (n, n) univerzális kód minimum Hamming távolsága $d = 1$, és az ilyen távolságban lévő kódszavak száma egy tetszőleges kódszótól $N_1 = n$.

Az $(n, n-1)$ SPC kód minimális Hamming súlya és távolsága $d = 2$, és az ilyen távolságban lévő kódszavak száma egy tetszőleges kódszótól $N_2 = n(n-1)/2$.

Az $(n, 1)$ ismétléses kód esetében a minimális Hamming távolság $d = n$, és az ilyen távolságban lévő kódszavak száma egy tetszőleges kódszótól $N_n = 1$.

Konvenció szerint az $(n, 0)$ triviális $C = O$ kód esetében a d értékét végtelennek szoktuk tekinteni.

Skaláris szorzat és ortogonalitás

Az $(\mathbb{F}_2)^n$ vektortérben a szimmetrikus bilineáris skaláris szorzatot az alábbi összefüggéssel definiáljuk:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} * \mathbf{y} = \mathbf{xy}^T = \sum_{i=1}^n x_i y_i,$$

ahol az \mathbb{F}_2 -beli elemekből álló n hosszúságú bináris sorozatoknak sorvektorok felelnek meg, és " T " a transzponáltat jelenti. Két sorozat akkor ortogonális, ha $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

Viszont az így definiált \mathbb{F}_2 skaláris szorzat nem rendelkezik a korábbi szigorúan pozitív tulajdonsággal, ugyanis a $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ összefüggésből nem következik az, hogy $\mathbf{x} = \mathbf{0}$, csak az, hogy \mathbf{x} éppen páros számú 1-et tartalmaz. Amiből az következik, hogy egy nem $\mathbf{x} \neq \mathbf{0}$ vektor ortogonális lehet önmagára. Mindebből látszik, hogy a $\langle \mathbf{x}, \mathbf{x} \rangle$ skaláris szorzat eltér az euklideszi térben megszokott normanégyzettől, és nem is lehet analóg módon kezelni azzal. Ebből is jól látszik, hogy a Hamming geometria alapvetően különbözik az euklideszi geometriától.

Részletezve: a Hamming térben a vetületek nem értelmezhetőek, ezért általában nem lehet megtalálni egy C bináris vektortér ortonormált \mathcal{G} bázisát.

Példa (Hamming távolságok)

Tekintsük a $(3, 2)$ SPC kódot, amely négy darab háromhosszúságú sorozatból áll:

$C = \{000, 011, 101, 110\}$. Ebből bármelyik két nem nulla kódszó bázisa a C kódnak, ugyanakkor ezek a kódszavak biztosan nem ortogonálisak egymásra.

Az (n, k) C kód C^\perp ortogonális kódja (duális kódja) azoknak az n hosszúságú bináris sorozatoknak a halmaza, melyek minden eleme ortogonális C minden elemére, azaz

$$C^\perp = \{\mathbf{y} \in \mathbb{F}_2^n \mid \langle \mathbf{y}, \mathbf{x} \rangle = 0, \text{ minden } \mathbf{x} \in C\}.$$

Nézzük meg ezután a C^\perp néhány elemi tulajdonságát:

- C^\perp egy $(n, n-k)$ lineáris kód, a kód \mathcal{H} bázisának a mérete $n-k$,

- Ha \mathcal{G} a C kód egy k elemű bázisa, akkor egy $n - k$ lineárisan független C^\perp -beli n hosszúságú sorozatból álló \mathcal{H} halmaz akkor és csak akkor bázisa a C^\perp duális kódnak, ha \mathcal{H} minden eleme ortogonális \mathcal{G} minden elemére,
- $(C^\perp)^\perp = C$.

A C kód \mathcal{G} bázisa k lineárisan független C -beli n hosszúságú bináris sorozatból áll, amit általában egy $k \times n$ méretű k rangú \mathbf{G} generátor mátrix formájában szoktunk ábrázolni. A C kód ilyenkor az összes $C = \{\mathbf{aG}, \mathbf{a} \in (\mathbb{F}_2)^k\}$ lineáris kombinációból áll. A C^\perp duális kód \mathcal{H} bázisa $n - k$ lineárisan független C^\perp -beli n hosszúságú bináris sorozatból áll, amihez egy $(n - k) \times n$ méretű $n - k$ rangú \mathbf{H} generátor mátrix tartozik, és $C^\perp = \{\mathbf{bH}, \mathbf{b} \in (\mathbb{F}_2)^{n-k}\}$. A fentebb említett második tulajdonság alapján kimondhatjuk, hogy C és C^\perp akkor és csak akkor duális kód pár, ha a generátor mátrixok teljesítik a $\mathbf{GH}^T = \mathbf{0}$ feltételt. A \mathbf{H} generátor mátrix \mathbf{H}^T transzponáltját a C kód paritás ellenőrző mátrixának nevezzük, ami azzal a tulajdonsággal rendelkezik, hogy egy $\mathbf{x} \in (\mathbb{F}_2)^n$ vektor akkor és csak akkor eleme a C kódnak, ha $\mathbf{xH}^T = \mathbf{0}$, mivel ilyenkor az adott \mathbf{x} akkor és csak akkor eleme a C^\perp kód duálisának, C -nek, ha ortogonális a C^\perp kód minden elemére.

Példa (Egyszerű kódok duálisai)

Az (n, n) univerzális és az $(n, 0)$ triviális kód duális párt képez. Az $(n, 1)$ ismétléses kód duális párja az $(n, n - 1)$ SPC kód. Megjegyezzük, hogy a $(2, 1)$ kód $\{00, 11\}$ egyszerre ismétléses és SPC kód, és önmaga duálisa is. Az ilyen kódokat önduális kódoknak nevezzük (az önduális tulajdonság az euklideszi vagy a komplex térben nem értelmezhető).

7.3. A lineáris blokk kódok az euklideszi térben

Ebben a fejezetben a bináris lineáris blokk kódok $s(C)$ euklideszi konstellációjának az alapvető paramétereit határozzuk meg az eredeti C bináris blokk kód paramétereinek a felhasználásával, azaz egy (n, k, d) kód esetében megadjuk N_d -t, a d Hamming súlyú C -beli kódszavak számát.

Legyen $s(C)$ dimenziója $N = n$ és a mérete $M = 2^k$. Ez azt jelenti, hogy egy blokkhoz (egy kódszóhoz) éppen k bit tartozik. A kód bitsebessége (névleges spektrális hatékonysága) $\rho = 2k/n$ [bit/két dimenzió]. Mivel $k \leq n$, ezért $\rho \leq 2$ [bit/két dimenzió], így a teljesítménykorlátozott tartományban vagyunk.

Az $s(C)$ vetület minden pontja megadható az $(\pm\alpha, \pm\alpha, \dots, \pm\alpha)$ formában, azaz minden ponthoz $n\alpha^2$ energia tartozik, tehát minden pont egy n dimenziós gömb felszínén fekszik, melynek a sugara $\sqrt{n}\alpha$. A kód blokkonkénti átlagos energiája $E(s(C)) = n\alpha^2$, és az átlagos bitenergia $E_b = n\alpha^2/k$. Ha két kódszó $(\mathbf{x}, \mathbf{y}) \in C$ Hamming távolsága $d_H(\mathbf{x}, \mathbf{y})$, akkor az $s(\mathbf{x})$ és $s(\mathbf{y})$ euklideszi vetületei $n - d_H(\mathbf{x}, \mathbf{y})$ helyen azonosak egymással, $d_H(\mathbf{x}, \mathbf{y})$ helyen pedig 2α a különbség közöttük. Ily módon, az euklideszi távolság négyzete

$$\|s(\mathbf{x}) - s(\mathbf{y})\|^2 = 4\alpha^2 d_H(\mathbf{x}, \mathbf{y}).$$

Sőt az n -dimenziós euklideszi térben $s(\mathbf{x})$ és $s(\mathbf{y})$ a skaláris szorzata

$$\langle s(\mathbf{x}), s(\mathbf{y}) \rangle = (n - d_H(\mathbf{x}, \mathbf{y}))\alpha^2 + d_H(\mathbf{x}, \mathbf{y})(-\alpha^2) = (n - 2d_H(\mathbf{x}, \mathbf{y}))\alpha^2.$$

Ezért $s(\mathbf{x})$ és $s(\mathbf{y})$ akkor és csak akkor ortogonális, ha $d_H(\mathbf{x}, \mathbf{y}) = n/2$. Emellett $s(\mathbf{x})$ és $s(\mathbf{y})$ akkor és csak akkor antipodális ($s(\mathbf{x}) = -s(\mathbf{y})$), ha $d_H(\mathbf{x}, \mathbf{y}) = n$.

Ezért a C kód minimális négyzetes euklideszi távolsága

$$d_{min}^2(s(C)) = 4\alpha^2 d_H(C) = 4\alpha^2 d,$$

ahol $d = d_H(C)$ a C kód minimális Hamming távolsága.

Ebből következően az $s(C)$ névleges kódolási nyeresége

$$\gamma_c(s(C)) = \frac{d_{\min}^2(s(C))}{4E_b} = \frac{kd}{n}.$$

Megállapíthatjuk tehát, hogy a $\gamma_c(s(C))$ értékét az (n, k, d) kód paramétereiből igen egyszerűen meg lehet határozni.

Emellett minden $s(\mathbf{x}) \in s(C)$ vektornak ugyanannyi $K_{\min}(s(\mathbf{x})) = N_d$ legközelebbi szomszédja, ahol N_d a legközelebbi szomszédok száma C -ben. Mivel $K_{\min}(s(C)) = N_d$, így $K_b(s(C)) = N_d/k$.

Ennek alapján az uniós korlát szerint a bit hibaarány a

$$P_b(e) \approx K_b(s(C)) Q \left(\sqrt{\gamma_c(s(C)) \frac{2E_b}{N_0}} \right) = \frac{N_d}{k} Q \left(\sqrt{\frac{kd}{n} \frac{2E_b}{N_0}} \right)$$

kifejezéssel közelíthető.

Összefoglalva az $s(C)$ bináris konstelláció paramétereit és teljesítőképességét az (n, k, d) paraméterek és a C kódhoz tartozó N_d érték segítségével igen egyszerűen meg lehet határozni.

Példák (Kódok tulajdonságai)

- Legyen C egy (n, k, d) lineáris bináris blokk kód, ahol d páratlan. Egészítsük ki az összes $\mathbf{x} \in C$ kódszót egy $p_x = \sum_i x_i$ paritás bittel, azaz minden új kódszó legyen $\mathbf{x}' = \{\mathbf{x}, p_x\}$ alakú.

1.) Az így keletkezett új C' kód nyilvánvalóan $n+1$ hosszúságú kódszavakból áll, és lineáris, mivel tartalmazza a nulla vektort, és két $\mathbf{x}', \mathbf{y}' \in C'$ új kódszó összege is kódszó, ugyanis $\mathbf{x}' + \mathbf{y}' = \{\mathbf{x}, p_x\} + \{\mathbf{y}, p_y\} = \{\mathbf{x} + \mathbf{y}, p_x + p_y\}$, és $\mathbf{x} + \mathbf{y} \in C$ az eredeti kód eleme, valamint $p_x + p_y$ ennek a paritás bitje. Tehát a C' kódra érvényes a csoport tulajdonság.

2.) Az új C' kód d' minimális Hamming távolsága biztosan $d+1$, ugyanis az eredeti kód esetén a minimális Hamming távolságú kódszavak közül az egyik Hamming súlya páros, a másiké páratlan, mivel a Hamming távolságuk páratlan. Így biztos hogy két minimális távolságú kódszó paritás bitjei nem lehetnek egyformák. Ebből az következik, hogy az új kódban az ezekből származtatott új kódszavak távolsága $d+1$ lesz. Emellett tudjuk, hogy a paritás bitek alkalmazásával az új kódban a kódszavak távolsága biztosan nem csökkenhet, mivel $d_H(\mathbf{x}', \mathbf{y}') = d_H(\{\mathbf{x}, p_x\}, \{\mathbf{y}, p_y\}) \geq d_H(\mathbf{x}, \mathbf{y})$. Ennek alapján az új C' egy $(n+1, k, d+1)$ paraméterű bináris lineáris blokk kód.

3.) Az új kód $\gamma_c(s(C'))$ névleges kódolási nyeresége biztosan nagyobb, mint az eredeti kódé, mivel:

$$\gamma_c(s(C)) = \frac{kd}{n} < \frac{k(d+1)}{n+1} = \gamma_c(s(C')), \quad \text{ha } n > d \text{ és } k > 1,$$

ami nyilvánvaló.

- Legyen C bináris lineáris blokk kód, ahol az $\mathbf{x} \in C$ kódszavak egy adott komponense vagy minden kódszóban 0, vagy a kódszavak felénél 0, másik felénél pedig 1. Példa erre az alábbi kód:

```
0 0 0 0 0,
0 1 0 0 0,
0 0 0 1 0,
0 1 0 1 0,
0 0 0 0 1,
0 1 0 0 1,
0 0 0 1 1,
0 1 0 1 1.
```

Erről a kódról belátható, hogy a kódszavak mindegyikéből elhagyhatók azok a komponensek, amelyek minden kódszóban 0 értékűek anélkül, hogy a kód teljesítőképessége romlana.

- 1.) Az így keletkezett C' kód paraméterei a következők lesznek: $k' = k$, mivel a kódszavak száma változatlan, $n' < n$, mivel a kódszavak hossza a minden kódszóból elhagyott komponensek számával csökken, $d' = d$, mivel a közös 0 komponensek elhagyása a kódszavak Hamming távolságát nem befolyásolja. Ennek alapján az új kód névleges kódolási nyeresége $\gamma_c(s(C'))$ nem csökkenhet az eredeti kódéhoz viszonyítva.
- 2.) Az új kód blokkonkénti átlagos energiája $E(s(C')) = n'\alpha^2$ kisebb, mint az eredeti kódé.
- 3.) Ha az eredeti kód kódszavainak nincs közös 0 komponense, akkor a kód $s(C)$ euklideszi vetületének az átlagos értéke $m(s(C)) = \mathbf{0}$, mivel az euklideszi vetület minden koordinátájában azonos számú $+1$ és -1 található. Ez utóbbi állítás a lineáris blokk kódokat előállító

$$C(\mathcal{G}) = \left\{ \sum_j a_j \mathbf{g}_j, a_j \in \mathbb{F}_2, 1 \leq j \leq k \right\}$$

általános kifejezésből közvetlenül következik, mivel a kódszavak i -dik összetevője felírható az

$$x_i = \left\{ \sum_j a_j g_{ji}, a_j \in \mathbb{F}_2, 1 \leq j \leq k \right\}$$

alakban, ahol g_{ji} a j -dik generátor i -dik eleme, és a_j felveszi a $(0, 1)$ halmaz elemeit. Ebből nyilvánvaló, hogy amennyiben legalább egy $g_{ji} \neq 0$, akkor a különböző $\mathbf{x} \in C$ kódszavak x_i komponensei azonos számban veszik fel a 0 és 1 értéket.

7.4. Reed-Muller kódok

A Reed-Muller (RM) kódok a bináris lineáris blokk kódok egy speciális végtelen elemű halmazát képezik. $n \leq 32$ hosszúságig ezek a kódok az eddig ismert legjobb kódok a d minimális Hamming távolság szempontjából, ha d kettő hatványa. Nagyobb kódhosszúság esetén általában az RM kódok nem az ismert legjobb kódok, de teljesítőképességük kiváló, ha a dekódolási komplexitást és a dekódoláshoz szükséges elegendően egyszerű ML algoritmusokat is figyelembe vesszük.

Minden $m \geq 0$ és $0 \leq r \leq m$ egész számokhoz tartozik egy RM kód, amit $\text{RM}(r, m)$ -mel jelölünk, és aminek a hosszúsága $n = 2^m$, minimális Hamming távolsága $d = 2^{m-r}$, $0 \leq r \leq m$.

$r = m$ esetén, $\text{RM}(m, m)$ definiíciószerűen a $(2^m, 2^m, 1)$ univerzális kód. A további vizsgálatokhoz szükséges az $r = -1$ értékhez tartozó RM kódokat definiálni, miszerint $\text{RM}(-1, m) = (2^m, 0, \infty)$, azaz a 2^m hosszúságú triviális kód. Így $m = 0$ -hoz két egy hosszúságú RM kód rendelhető, az $(1, 1, 1)$ univerzális kód $\text{RM}(0, 0)$ és az $(1, 0, \infty)$ triviális kód $\text{RM}(-1, 0)$.

A többi RM kód $m \geq 0$ és $0 \leq r \leq m$ esetén ezekből az elemi kódokból származtatható a következő, eredetileg Plotkintól származó $|\mathbf{u} \mid \mathbf{u} + \mathbf{v} \mid$ speciális hossz-kétszerező algoritmus segítségével. Az $\text{RM}(r, m)$ kód az $\text{RM}(r-1, m-1)$ és az $\text{RM}(r, m-1)$ kódokból állítható elő az

$$\text{RM}(r, m) = \{(\mathbf{u}, \mathbf{u} + \mathbf{v}) \mid \mathbf{u} \in \text{RM}(r, m-1), \mathbf{v} \in \text{RM}(r-1, m-1)\}$$

rekurzív formula segítségével.

A fenti konstrukció alapján egyszerűen belátható, hogy az így előállított kódok a következő tulajdonságokkal bírnak:

- $\text{RM}(r, m)$ bináris lineáris kód $n = 2^m$ hosszúsággal és $k(r, m) = k(r, m-1) + k(r-1, m-1)$ dimenzióval,
- A kódcsalád egymásba ágyazódik, ugyanis $\text{RM}(r-1, m) \subseteq \text{RM}(r, m)$,

- Az $\text{RM}(r, m)$ kód minimális Hamming távolsága $d = 2^{m-r}$, ha $r \geq 0$ (ha $r = -1$, akkor $d = \infty$).

A fentiek alapján egyszerűen belátható, hogy ezek a tulajdonságok az $\text{RM}(0, 0)$ és az $\text{RM}(-1, 0)$ esetén teljesülnek. $m \geq 1$ esetén a linearitásra és a $\text{RM}(r, m)$ kód hosszúságára vonatkozó állítások a konstrukcióból közvetlenül származtathatók. A kód dimenziója (mérete) pedig abból a tényből következik, hogy $(\mathbf{u}, \mathbf{u} + \mathbf{v})$ akkor és csak akkor $\mathbf{0}$, ha $\mathbf{u} = \mathbf{v} = \mathbf{0}$.

Az alábbiakban igazoljuk, hogy az $\text{RM}(r, m)$ kódok méretére vonatkozó $k(r, m) = k(r, m-1) + k(r-1, m-1)$ rekurzió a

$$k(r, m) = \sum_{0 \leq j \leq r} \binom{m}{j}$$

zárt formulához vezet, ahol $\binom{m}{j} = \frac{m!}{j!(m-j)!}$ a szokásos kombinatorikus koefficiens. Ez egyszerűen belátható, mivel a rekurzió szerint

$$k(r, m) = \sum_{0 \leq j \leq r} \binom{m-1}{j} + \sum_{0 \leq j \leq r-1} \binom{m-1}{j} = 1 + \sum_{1 \leq j \leq r} \left(\binom{m-1}{j} + \binom{m-1}{j-1} \right) = \sum_{0 \leq j \leq r} \binom{m}{j},$$

mivel a binomiális együtthatókra ismert az

$$\begin{aligned} \binom{m-1}{j} + \binom{m-1}{j-1} &= \frac{(m-1)!}{j!(m-1-j)!} + \frac{(m-1)!}{(j-1)!(m-1-j+1)!} = \\ &= \frac{(m-1)!}{(j-1)!(m-1-j)!} \left(\frac{1}{j} + \frac{1}{m-j} \right) = \frac{(m-1)!}{(j-1)!(m-1-j)!} \binom{m}{j(m-j)} = \binom{m}{j} \end{aligned}$$

azonosság, és $\binom{m-1}{0} = 1$.

A kódcsalád egymásba ágyazódási tulajdonsága az m -dik szinten egyenesen következik az $m-1$ -dik szinten érvényes állításból.

Az alábbiakban pedig belátjuk, hogy az $\text{RM}(r, m)$ kód minimális nem zérus Hamming súlya 2^{m-r} :

(a) Ha $\mathbf{u} = \mathbf{0}$, akkor $w_H((\mathbf{0}, \mathbf{v})) = w_H(\mathbf{v}) \geq 2^{m-r}$ ha $\mathbf{v} \neq \mathbf{0}$, mivel $\mathbf{v} \in \text{RM}(r-1, m-1)$.

(b) Ha $\mathbf{u} + \mathbf{v} = \mathbf{0}$, akkor $\mathbf{u} = \mathbf{v} \in \text{RM}(r-1, m-1)$ és $w_H((\mathbf{v}, \mathbf{0})) \geq 2^{m-r}$ ha $\mathbf{v} \neq \mathbf{0}$.

(c) Ha $\mathbf{u} \neq \mathbf{0}$, akkor mind \mathbf{u} , mind $\mathbf{u} + \mathbf{v}$ is eleme az $\text{RM}(r, m-1)$ (mivel $\text{RM}(r-1, m-1)$ részhalmaza az $\text{RM}(r, m-1)$ kódnak), ily módon

$$w_H((\mathbf{u}, \mathbf{u} + \mathbf{v})) = w_H(\mathbf{u}) + w_H(\mathbf{u} + \mathbf{v}) \geq 2 \times 2^{m-r-1} = 2^{m-r}.$$

Az azonosság pedig nyilvánvalóan fennáll $(\mathbf{0}, \mathbf{v})$, $(\mathbf{v}, \mathbf{0})$ vagy (\mathbf{u}, \mathbf{u}) esetén, ha \mathbf{v} vagy \mathbf{u} aktuális értékeként a saját kódjuk minimális Hamming súlyú vektorát választjuk.

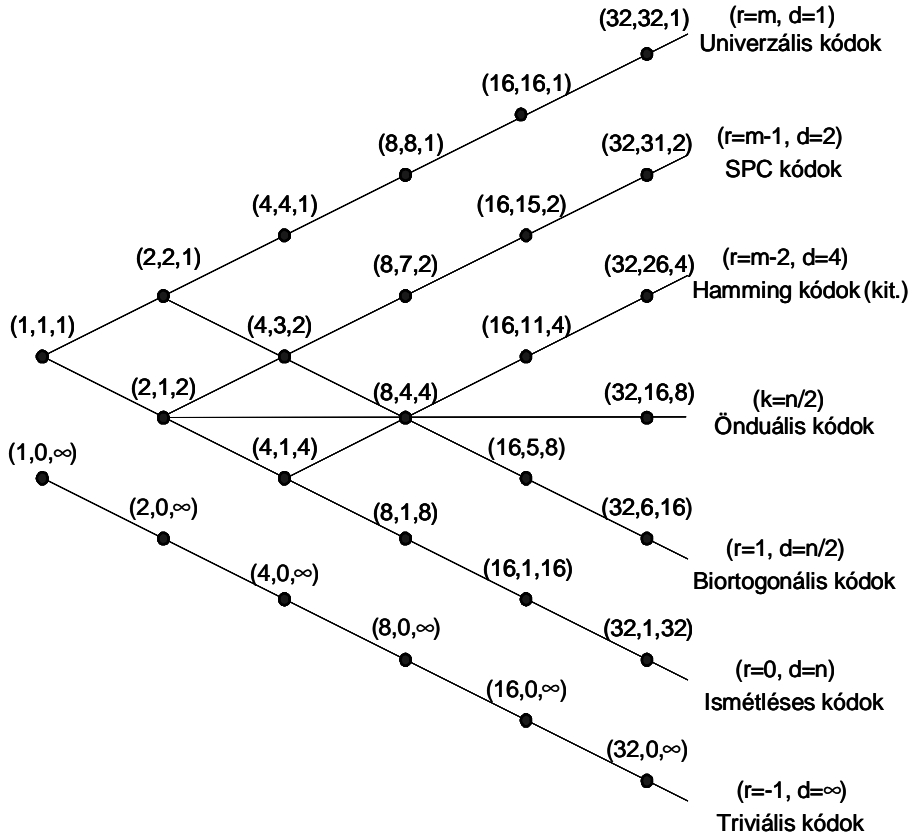
A 7.2. ábrán megadtuk a különböző Reed-Muller kódok egymásra épülését a $|\mathbf{u}| \mathbf{u} + \mathbf{v}|$ konstrukció szerint.

A táblázatban minden RM kód két másik kód között helyezkedik el a $|\mathbf{u}| \mathbf{u} + \mathbf{v}|$ konstrukciós szabálynak megfelelően. A táblázatból azonnal leolvasható az aktuális kód hossza és k dimenziója.

Példák (A Reed-Muller kódok néhány tulajdonsága)

- Számoljuk ki a 64 hosszúságú RM kódok (k, d) és N_d paramétereit. Ehhez használjuk fel a legközelebbi szomszédok számára vonatkozó ismert zárt formulát:

$$N_d = 2^r \prod_{0 \leq i \leq m-r-1} \frac{2^{m-i} - 1}{2^{m-r-i} - 1},$$



7.2. ábra. A Reed-Müller kódok táblázata

amely megadja az $RM(r, m)$ Reed-Muller kód $d = 2^{m-r}$ minimális Hamming súlyú kódszavainak a számát. Ezek alapján a 64 hosszúságú RM kódok paraméterei az alábbiak:

$m = 6,$	$r = 6,$	$n = 64,$	$d = 1,$	$k = 64,$	$N_d = 64$
$m = 6,$	$r = 5,$	$n = 64,$	$d = 2,$	$k = 63,$	$N_d = 2016$
$m = 6,$	$r = 4,$	$n = 64,$	$d = 4,$	$k = 57,$	$N_d = 10416$
$m = 6,$	$r = 3,$	$n = 64,$	$d = 8,$	$k = 42,$	$N_d = 11160$
$m = 6,$	$r = 2,$	$n = 64,$	$d = 16,$	$k = 22,$	$N_d = 2604$
$m = 6,$	$r = 1,$	$n = 64,$	$d = 32,$	$k = 7,$	$N_d = 126$
$m = 6,$	$r = 0,$	$n = 64,$	$d = 64,$	$k = 1,$	$N_d = 1$
$m = 6,$	$r = -1,$	$n = 64,$	$d = \infty,$	$k = 0,$	$N_d = 0$

- Az $r = m - 1$ RM kódok SPC (single parity check, paritás ellenőrző) kódok $d = 2$ minimális Hamming távolsággal. Ezeknek a kódoknak a névleges kódolási nyeresége $\gamma_c = 2k/n$, ami 2-höz tart (3,01 dB), ha $n \rightarrow \infty$, mivel $n = 2^m$ és $k = 2^m - 1$. Ugyanakkor $N_d = 2^m(2^m - 1)/2$, és $K_b = N_d/k = 2^{m-1}$, ami ∞ -hez tart, ha n minden határon túl nő. Ez erősen korlátozza a kód effektív kódolási nyereségét.

Határozzuk meg m optimális értékét, ahol az SPC kódok effektív kódolási nyeresége maximális. Használjuk ehhez a korábban megismert

$$\gamma_{eff}(\mathcal{A}) \approx \gamma_c(\mathcal{A}) - (0,2) \log_2(K_b(\mathcal{A})) \quad \text{dB-ben.}$$

közelítést, ami a $P_b = 10^{-5}$ hibaaarányhoz tartozik.

Átírva a dB-ben értelmezett kifejezést lineáris skálára, a

$$\gamma_{eff}(\mathcal{A}) \approx \frac{\gamma_c(\mathcal{A})}{10^{(0,2)\log_2(K_b(\mathcal{A}))/10}} = \gamma_c(\mathcal{A})10^{-(0,02)\log_2(K_b(\mathcal{A}))}$$

kifejezéshez jutunk, amit az SPC kódok esetében érvényes $K_b(\mathcal{A}) = 2^{m-1}$ és $\gamma_c(\mathcal{A}) = 2(2^m - 1)/2^m$ felhasználásával a

$$\gamma_{eff}(\mathcal{A}) \approx \frac{2(2^m - 1)}{2^m} 10^{-(0,02)(m-1)} = 2 \left(1 - e^{-m\ln(2)}\right) e^{-0,02\ln(10)(m-1)}$$

alakra lehet hozni. Ennek az m szerinti deriváltja az

$$2\ln(2)e^{-m\ln(2)}e^{-0,02\ln(10)(m-1)} - 2 \left(1 - e^{-m\ln(2)}\right) 0,02\ln(10)e^{-0,02\ln(10)(m-1)} = 0$$

helyen veszi fel a nulla értéket. Ebből egyszerű átalakítások után az

$$m = \frac{1}{\ln(2)} \ln \frac{\ln(2) + 0,02\ln(10)}{0,02\ln(10)} = 4,0046$$

optimális értékhez jutunk. Ez azt jelenti, hogy az SPC kódok kb. $m = 4$ értéknél biztosítanak maximális kódolási nyereséget, aminek az értéke kb. 1.63.

- Az $r = m - 2$ RM kódok EH (extended Hamming, kiegészített Hamming) kódok $d = 4$ minimális Hamming távolsággal. Ezeknek a kódoknak a névleges kódolási nyeresége $\gamma_c = 4k/n$, ami 4-hez tart (6,02 dB), ha $n \rightarrow \infty$, mivel $n = 2^m$ és $k = 2^m - m - 1$. Ugyanakkor $N_d = 2^m(2^m - 1)(2^m - 2)/24$, és K_b most is ∞ -hez tart, ha n minden határon túl nő.

Határozzuk meg m optimális értékét az EH kódok esetén is az előbbi analízishez hasonló eljárás alkalmazásával.

Felhasználva $\gamma_{eff}(\mathcal{A})$ közelítő értékét, és behelyettesítve az aktuális paramétereket, most a

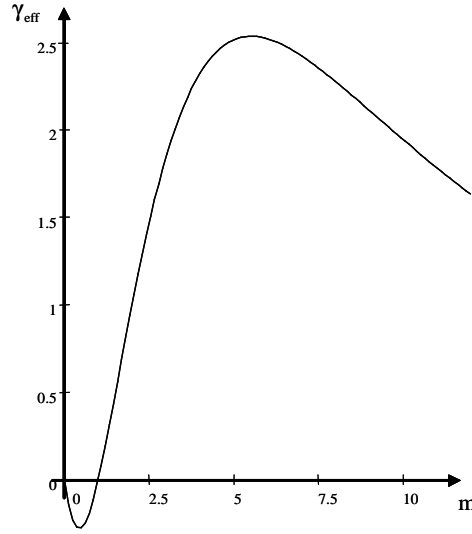
$$\gamma_{eff}(\mathcal{A}) \approx 4(2^m - m - 1)2^{-m}10^{-(0,02)\log_2\left(\frac{2^m(2^m-1)(2^m-2)}{24(2^m-m-1)}\right)}$$

kifejezést kapjuk. Ennek az m szerinti deriváltja zárt alakban bonyolultan kezelhető, ezért az optimumot grafikusán határozzuk meg a fenti függvény ábrázolásával (lásd a 7.3. ábrát).

Ez azt jelenti, hogy az EH kódok optimális effektív kódolási nyereségét kb. $m = 6$ értéknél kapjuk, melynek az értéke kb. 2.52. A görbe negatív értékeivel kapcsolatban megjegyezzük, hogy $m \leq 1$ -nél az EH kódok nem értelmezhetők.

- Az $r = 1$ paraméterű (első rendű Reed-Muller) kódok igen érdekesek, mivel 2^{m+1} méretű $n = 2^m$ dimenziós biortogonális jelteret generálnak, melynek a névleges kódolási nyeresége $(m + 1)/2 \rightarrow \infty$, ha m minden határon túl nő. Ismert, hogy ha $n \rightarrow \infty$, akkor az ilyen kódokkal tetszőlegesen kis $\Pr(e)$ hibaarány érhető el, bármilyen E_b/N_0 jel-zaj viszony esetén az abszolút Shannon határ, $E_b/N_0 > \ln 2$ (-1.59 dB) felett.
- A korábbiakból tudjuk, hogy az elsőrendű RM(1, m) Reed-Muller kód paraméterei ($n = 2^m, k = m + 1, d = 2^{m-1}$), és az RM(0, m) ($n = 2^m, k = 1, d = 2^m$) paraméterű ismétléses kód ennek részhalmaza (alkódja).
- Az alábbiakban megmutatjuk, hogy az RM(1, m) kód egyetlen nulla Hamming súlyú, egyetlen 2^m Hamming súlyú és $2^{m+1} - 2$ számú 2^{m-1} Hamming súlyú kódszóból áll. A speciális hosszszerező algoritmus alapján

$$\text{RM}(1, m) = \{(\mathbf{u}, \mathbf{u} + \mathbf{v}) \mid \mathbf{u} \in \text{RM}(1, m - 1), \mathbf{v} \in \text{RM}(0, m - 1)\},$$



7.3. ábra. Az EH kódok effektív kódolási nyeresége az m függvényében

és az $\text{RM}(0, m-1) = \{\mathbf{0}, \mathbf{1}\}$ kód két elemet tartalmaz, az 2^{m-1} hosszúságú csupa 0 és az csupa 1 vektort.

Tudjuk, hogy az $\text{RM}(1, 1) = \{(00), (01), (10), (11)\}$ kód tartalmazza minden kódszó inverzét. A hossz-kétszerező algoritmus alapján

$$\text{RM}(1, m) = \{(\mathbf{u}, \mathbf{u} + \mathbf{1}), (\mathbf{u}, \mathbf{u} + \mathbf{0}) \mid \mathbf{u} \in \text{RM}(1, m-1), \mathbf{v} \in \text{RM}(0, m-1) = \{\mathbf{0}, \mathbf{1}\}\},$$

és, ha $\text{RM}(1, m-1)$ tartalmazza minden kódszó inverzét, akkor $\text{RM}(1, m)$ is tartalmazza, mivel $(\mathbf{u}, \mathbf{u} + \mathbf{1})$ inverze $(\mathbf{u} + \mathbf{1}, \mathbf{u})$, és (\mathbf{u}, \mathbf{u}) inverze $(\mathbf{u} + \mathbf{1}, \mathbf{u} + \mathbf{1})$. Ennek alapján megállapíthatjuk, hogy a $\mathcal{C} = \text{RM}(1, m)$ kód 2^m számú inverz kódszó párból áll, azaz tartalmazza az \mathbf{x} és $\mathbf{x} + \mathbf{1}$ kódszavakat is ha $\mathbf{x} \in \mathcal{C}$.

Visszatérve az eredeti problémához, miszerint az $\text{RM}(1, m)$ kód egy nulla Hamming súlyú, egy 2^m Hamming súlyú és $2^{m+1} - 2$ számú 2^{m-1} Hamming súlyú kódszóból áll, megállapíthatjuk, hogy az $\text{RM}(1, 1)$ kódra az állítás igaz, és a hossz-kétszerező algoritmus alapján egy $\mathbf{x} \in \mathcal{C}$ vektor Hamming súlya $w_H(\mathbf{x}) = 2w_H(\mathbf{u}) = 2 \cdot 2^{m-2} = 2^{m-1}$ vagy

$$w_H(\mathbf{x}) = w_H(\mathbf{u}) + w_H(\mathbf{u} + \mathbf{1}) = 2^{m-2} + (2^{m-1} - 2^{m-2}) = 2^{m-1},$$

ha $\mathbf{u} \in \text{RM}(1, m-1)$ és $\mathbf{u} \neq \mathbf{1}, \mathbf{0}$. Ha $\mathbf{u} = \mathbf{1}$, akkor $w_H(\mathbf{u}, \mathbf{u}) = 2^m$, ha $\mathbf{u} = \mathbf{0}$, akkor $w_H(\mathbf{u}, \mathbf{u}) = 0$. Ennek alapján az állítás igaz.

- A következőkben megmutatjuk, hogy az az $\text{RM}(1, m)$ kód euklideszi térbeli vetülete egy $M = 2^{m+1}$ méretű biortogonális jelkészlet.

A korábban vizsgált $\text{RM}(1, 1)$ kód euklideszi térbeli vetülete triviálisan biortogonális az $N = 2^1 = 2$ dimenziós térben. A hossz-kétszerező algoritmus alapján nyilvánvaló, hogy

$$\langle s(\mathbf{u}, \mathbf{u}), s(\mathbf{u}, \mathbf{u} + \mathbf{1}) \rangle = \langle s(\mathbf{u}), s(\mathbf{u}) \rangle + \langle s(\mathbf{u}), s(\mathbf{u} + \mathbf{1}) \rangle = \langle s(\mathbf{u}), s(\mathbf{u}) \rangle - \langle s(\mathbf{u}), s(\mathbf{u}) \rangle = 0,$$

azaz az $s(\mathbf{u}, \mathbf{u})$ és $s(\mathbf{u}, \mathbf{u} + \mathbf{1})$ vektorok merőlegesek egymásra, és az

$$\langle s(\mathbf{u}, \mathbf{u}), s(\mathbf{u}', \mathbf{u}') \rangle = \langle s(\mathbf{u}), s(\mathbf{u}') \rangle + \langle s(\mathbf{u}), s(\mathbf{u}') \rangle = 2\langle s(\mathbf{u}), s(\mathbf{u}') \rangle$$

összefüggés alapján az $s(\mathbf{u}, \mathbf{u})$ és $s(\mathbf{u}', \mathbf{u}')$ merőlegesek, ha az $s(\mathbf{u})$ és $s(\mathbf{u}')$ merőlegesek egymásra, és egymás mínusz egyszeresei, ha $s(\mathbf{u}) = -s(\mathbf{u}')$. Teljes indukcióval ezután belátható, hogy az $s(\mathcal{C})$ jeltér 2^m dimenziós, és 2^m inverz jelpárt tartalmaz, ahol a jelpárok ortogonálisak egymásra.

- Az $RM(1, m)$ kódból válasszuk ki azokat a kódszavakat, amelyeknek egy adott helyértékű (például a j -edik) eleme 0 értékű. Lássuk be, hogy az így létrehozott C' kód egy $(n = 2^m, k = m, d = 2^{m-1})$ paraméterű bináris lineáris kód, melynek az euklideszi térben egy $M = 2^m$ méretű ortogonális jelkészlet felel meg.

A C' kód lineáris, mivel elemei az eredeti lineáris C kód elemei, ami az összeadásra zárt, azaz ha $\mathbf{x}, \mathbf{y} \in C$, akkor $\mathbf{x} + \mathbf{y} \in C$. Ugyanakkor tudjuk, hogy ha $\mathbf{x}, \mathbf{y} \in C'$, akkor $\mathbf{x} + \mathbf{y} \in C'$, tehát a C' kód lineáris, emellett biztosan eleme a $\mathbf{0}$ vektor.

Nyilvánvaló, hogy $C' \subseteq C$, ezért biztosan igaz, hogy $n = 2^m$. A kódszavak ilyen szelekciója esetén ha $\mathbf{x} \in C'$ akkor \mathbf{x} inverze nem lehet eleme C' -nek, ugyanis ha \mathbf{x} egy adott helyértékű összetevője 0, akkor azon a helyértéken \mathbf{x} inverzének az összetevője 1. Ez azt jelenti, hogy a kódszelekcióval elhagyjuk minden kiválasztott kódszó inverzét, azaz a kódszavak száma felére csökken, tehát $k = m$.

Mivel az $RM(1, m)$ kód euklideszi térbeli vetülete egy $M = 2^{m+1}$ méretű biortogonális jelkészlet, az inverz párok elhagyásával a megmaradó kódszavak vetületei az euklideszi térben ortogonális jelkészletet alkotnak, és a minimális Hamming távolság nem változik.

- A C' kód minden kódszavának töröljük ki a közös 0 értékű (például a j -edik helyértékű) elemét. Lássuk be, hogy az így létrehozott C'' kód egy $(n = 2^m - 1, k = m, d = 2^{m-1})$ paraméterű bináris lineáris kód, melynek az euklideszi térben egy $M = 2^m$ méretű szimplex jelkészlet felel meg.

A C'' kód elemei az eredeti lineáris C' kód elemeiből származnak egyetlen közös 0 értékű (például a j -edik helyértékű) elem törlésével. Tudjuk, hogy a C' kód az összeadásra zárt, azaz ha $\mathbf{x}, \mathbf{y} \in C'$, akkor $\mathbf{x} + \mathbf{y} \in C'$. Ezért igaz, hogy a C'' kód lineáris, ugyanis ha $\mathbf{x}, \mathbf{y} \in C''$, akkor $\mathbf{x} + \mathbf{y} \in C''$, mivel az összeadásnál a törölt elemet nem kell figyelembe venni, emellett C'' -nek biztosan eleme a $\mathbf{0}$ vektor.

Nyilvánvaló, hogy minden kódszó egy elemének elhagyásával $n = 2^m - 1$, és a C'' kódszavainak a száma azonos a C' kódszavainak a számával.

Emellett az $RM(1, m)$ kód különböző kódszavainak azonos sorszámú összetevője azonos számban veszi fel a logikai 0 és logikai 1 értéket. Ez teljesül az $RM(1, 1) = \{(00), (01), (10), (11)\}$ kód esetére, és teljesül például az $RM(0, 1) = \{(00), (11)\}$ kód esetére is. Teljes indukcióval könnyen belátható, hogy ha $\mathbf{u} \in C'$, és C' a $C = RM(1, m)$ kódból származtatott ortogonális euklideszi vetülettel rendelkező kód, akkor a nulla elemeket tartalmazó (például j -edik) összetevőn kívül minden más összetevő azonos számban veszi fel a logikai 0 és logikai 1 értéket. Ezért a hossz-kétszerező algoritmus alapján az $(\mathbf{u}, \mathbf{u} + \mathbf{1})$ és a $(\mathbf{u}, \mathbf{u} + \mathbf{0})$ vektorok j -edik összetevőjén kívül minden más összetevő azonos számban vesz fel a logikai 0 és logikai 1 értéket. Ezért igaz, hogy

$$\mathbf{m}(C') = \frac{1}{|C'|} \sum_{i=1}^M s(\mathbf{x}_i) = \frac{1}{|C'|} \sum_{i=1}^M \{s(x_{i_1}), \dots, s(x_{i_j}), \dots, s(x_{i_m})\} = \{0, \dots, s(x_{i_j}) = 1, \dots, 0\}.$$

Ebből pedig egyenesen következik, hogy a kódszavak j -edik összetevőjének törlése esetén a C'' kód kódszavai előállíthatók az $\mathbf{x}'' = \mathbf{x}' - \mathbf{m}(C')$ formában, ami a szimplex jelkészlet definíciója.

- Az $RM(r, m)$ bináris lineáris kód duál kódja az $RM(m - r - 1, m)$ Reed-Muller kód. Ezért az ismétléses és az SPC kódok, illetve a biortogonális és a kiegészített Hamming kódok duális párokat alkotnak.
- Az $k = n/2$ paraméterű RM kódok önduálisak. Az ilyen kódok hossza $n = 2^m$ és $r = (m - 1)/2$, $k = 2^{m-1}$ és $d = 2^{(m+1)/2}$, ezért a kódok névleges kódolási nyeresége $\gamma_c = kd/n = 2^{(m-1)/2}$ (m páratlan), ami m növelésével minden határon túl nő. Ebből az várható, hogy ezzel a kóddal

tetszőlegesen kis $\Pr(e)$ érhető el ha E_b/N_0 kis mértékben meghaladja a Shannon határt. $\rho = 1$ [bit/két dimenzió] esetén ez > 1 (0 dB).

A Reed-Muller kódok effektív kódolási nyeresége

Az alábbi táblázatban megadjuk néhány Reed-Muller kód jellegzetes paramétereit, a ρ névleges spektrális hatékonyságot, a γ_c névleges kódolási nyereséget, N_d -t, a legközelebbi szomszédok átlagos számát, a K_b faktort és γ_{eff} -et, a becsült effektív kódolási nyereséget $P_b(e) = 10^{-5}$ értéknél.

Kód	ρ	γ_c (dB)	N_d	K_b	γ_{eff} (dB)	
(8, 7, 2)	1,75	7/4	2,43	28	4	2,0
(8, 4, 4)	1,00	2	3,01	14	4	2,6
(16, 15, 2)	1,88	15/8	2,73	120	8	2,1
(16, 11, 4)	1,38	11/4	4,39	140	13	3,7
(16, 5, 8)	0,63	5/2	3,98	30	6	3,5
(32, 31, 2)	1,94	31/16	2,87	496	16	2,1
(32, 26, 4)	1,63	13/4	5,12	1240	48	4,0
(32, 16, 8)	1,00	4	6,02	620	39	4,9
(32, 6, 16)	0,37	3	4,77	62	10	4,2
(64, 63, 2)	1,97	63/32	2,94	2016	32	1,9
(64, 57, 4)	1,78	57/16	5,52	10416	183	4,0
(64, 42, 8)	1,31	21/4	7,20	11160	266	5,6
(64, 22, 16)	0,69	11/2	7,40	2604	118	6,0
(64, 7, 32)	0,22	7/2	5,44	126	18	4,6

7.5. A bináris blokk kódok dekódolása

Ebben a fejezetben először megmutatjuk, hogy bináris blokk kódok minimális távolságú dekódolása (MD) ekvivalens a maximális megbízhatóságú dekódolással. Azután bemutatjuk, hogy milyen hátrányokkal jár a kemény döntés (hard decision), és megmutatjuk a klasszikus hibajavítás hatását a rendszer teljesítőképességére. Ezután azt vizsgáljuk, hogy ezt a veszteséget hogyan lehet csökkenteni a törlés alkalmazásával, illetve az általánosított minimális távolságú dekódolással.

A maximális megbízhatóságú dekódolás

Eddig minden teljesítőképesség becslésnél azt feltételeztük, hogy a minimális távolságú dekódolási szabályt (MD) alkalmaztuk. Másképpen megfogalmazva feltételeztük, hogy egy adott $\mathbf{r} \in \mathbb{R}^n$ bemeneti sorozat esetén a vevő megtalálja azt az $s(\mathbf{x})$, $\mathbf{x} \in \mathcal{C}$ jelet, amelynek az $\|\mathbf{r} - s(\mathbf{x})\|^2$ euklideszi távolság négyzete minimális. Megmutatjuk, hogy bináris kódok esetén az MD dekódolás ekvivalens a maximális megbízhatóságú dekódolással. Mivel $\|s(\mathbf{x})\|^2 = n\alpha^2$, függetlenül a \mathbf{x} tényleges választásától, az MD dekódolás azonos a maximális skaláris szorzatú dekódolással: miszerint meg kell találni azt az $s(\mathbf{x})$ $\mathbf{x} \in \mathcal{C}$ jelet, amelynél az

$$\langle \mathbf{r}, s(\mathbf{x}) \rangle = \sum_k r_k s(x_k)$$

skaláris szorzat maximális. Mivel $s(x_k) = (-1)^{x_k} \alpha$, ezért a skaláris szorzat az

$$\langle \mathbf{r}, s(\mathbf{x}) \rangle = \alpha \sum_k r_k (-1)^{x_k} = \alpha \sum_k |r_k| \operatorname{sgn}(r_k) (-1)^{x_k}$$

alakban adható meg. Az $\operatorname{sgn}(r_k) \in \{\pm 1\}$ nem más, mint az r_k előjele, amit általában kemény döntésként (hard decision) szoktunk emlegetni, és ami azt mutatja meg, hogy a két lehetséges jel közül

melyik valószínűbb az adott helyértéken, függetlenül a többi koordinátától. Az $|r_k|$ nagysága pedig a kemény döntés megbízhatóságát méri. Az ilyen döntést tehát úgy is interpretálhatjuk, hogy ezzel meg lehet találni azt az $\mathbf{x} \in C$ kódszót, mely maximalizálja az

$$r(\mathbf{x}|\mathbf{r}) = \sum_k |r_k| (-1)^{e(x_k, r_k)}$$

megbízhatóság értékét, ahol az $e(x_k, r_k)$ "hiba" 0, ha $s(x_k)$ és r_k előjele azonos, és 1, ha különböző. Éppen ezért ezt az eljárást maximális megbízhatóságú dekódolásnak nevezzük. Bármelyik optimális dekódolási eljárást könnyen meg lehet valósítani kis méretű $s(C)$ jelterek esetében.

Igaz azonban, hogy speciális ötletek nélkül az eljárások során minden $\mathbf{x} \in C$ kódszó esetén ki kell számolni a skaláris szorzat vagy a megbízhatóság értékét, ezért a módszer a gyakorlatban nehezen használható, ha a kódszavak 2^k száma nagy. Egyszerűbb, kis komplexitású és elegendően pontos dekódolási algoritmusok kifejlesztése a gyakorlati kódolással kapcsolatos kutatások egyik legfontosabb témája.

Például vizsgáljuk meg a Wagner féle dekódolási módszert, ami optimális az $(n, n-1, 2)$ SPC kódok esetében, és 2^{n-1} -nél sokkal kevesebb lépést igényel.

Példa: (A Wagner dekódolási szabály)

Legyen C egy $(n, n-1, 2)$ SPC kód. Ezekre a kódokra a Wagner dekódolási szabály a következő: végezzük el a kemény döntést minden r_k szimbólumra, és ellenőrizzük, vajon az kemény döntés eredményeképpen keletkező bináris vektor eleme-e a C -nek. Ha igen, akkor a dekódolást befejeztük, ha nem, akkor változtassuk meg a legkisebb $|r_k|$ megbízhatóságú szimbólum előjelét. Egyszerűen belátható, hogy az így kapott bináris vektor biztosan eleme C -nek, mivel az $(n, n-1, 2)$ SPC kód az összes páros Hamming súlyú kódszóból áll, és kódszavak közötti minimális Hamming távolság 2, ennek alapján, ha egy kódszó nem tartozik a C -be, akkor biztosan páratlan a Hamming súlya, ami egy helyérték előjelének megváltoztatásával párossá válik. Ezért az új kódszó biztosan eleme lesz az $(n, n-1, 2)$ SPC kódnak, ugyanakkor az ehhez a kódszóhoz tartozó $r(\mathbf{x}|\mathbf{r})$ megbízhatóság maximális. Ebből következik, hogy a Wagner dekódolási szabály az SPC kódok esetében optimális.

Kemény döntés és hibajavítás

A bináris blokk kódok dekódolásával foglalkozó korai megoldásokban mindig kemény döntést alkalmaztak minden szimbólumra, azaz a döntés eredményeként egy n hosszúságú $\mathbf{y} \in (\mathbb{F}_2)^n$ bináris sorozatot kaptak, és a dekódolás során megkeresték azt az $\mathbf{x} \in C$ kódszót, aminek a Hamming távolsága az \mathbf{y} -tól minimális. Ezt az eljárást nevezzük hibajavításnak.

Ha C egy lineáris (n, k, d) kód, akkor, mivel a Hamming metrika valódi metrika, nem keletkezhet hiba, ha az \mathbf{x} kódszó dekódolása során keletkező kemény döntési hibák száma $t = d_H(\mathbf{x}, \mathbf{y})$ kisebb, mint a kódra jellemző minimális Hamming távolság fele, $t < d/2$. A blokk kódok több osztályára léteznek olyan hatékony algebrai hibajavító algoritmusok, melyek garantálják azt, hogy hibátlanul dekódolják az üzenetet, ha $2t < d$.

Példa (Hamming kódok)

Az elsőnek alkalmazott bináris hibajavító kód a Hamming kód volt. Egy C $(2^m - 1, 2^m - m - 1, 3)$ Hamming kód a $(2^m, 2^m - m - 1, 4)$ paraméterű $RM(m-2, m)$ kiegészített Hamming kódból származtatható az egyik koordináta törlésével. A Hamming kód duális kódja a C^\perp $(2^m - 1, m, 2^{m-1})$ kód, melynek az euklideszi térben egy 2^m szimplex jelkonstelláció felel meg.

Példaképpen tekintsük a legegyszerűbb Hamming kódot az $(3, 1, 3)$ paraméterű ismétléses kódot, melynek a duálisa a $(3, 2, 2)$ SPC kód, melynek az euklideszi térben a 4 szimplex jelkonstelláció felel meg.

A C^\perp generátor mátrixa egy $m \times (2^m - 1)$ \mathbf{H} mátrix, amelynek $2^m - 1$ oszlopa valamilyen sorrendben felveszi az összes nem nulla értékű m hosszúságú bináris sorozatot (egyébként nem lenne

képes bármely egyszeres hiba javítására). Mivel $d = 3$, a Hamming kód képes minden egyszeres hiba javítására. Erre egy igen egyszerű módszert lehet adni. Először számítsuk ki a szindrómát

$$\mathbf{yH}^T = (\mathbf{x} + \mathbf{e})\mathbf{H}^T = \mathbf{eH}^T$$

ahol $\mathbf{e} = \mathbf{x} + \mathbf{y}$. Ha $\mathbf{yH}^T = \mathbf{0}$, akkor $\mathbf{y} \in C$ -ről feltételezzük, hogy korrekt. Ha $\mathbf{yH}^T \neq \mathbf{0}$, akkor az \mathbf{yH}^T szindróma a \mathbf{H}^T egy sora, és egyetlen hibát feltételezve a hiba az ehhez a sorhoz tartozó pozícióban történt meg. Egyébként tudjuk, hogy bármely $\mathbf{y} \in (\mathbb{F}_2)^n$ sorozatot kódszóvá lehet alakítani legfeljebb egyetlen bit megváltoztatásával.

Ez abból következik, hogy 2^{n-m} számú 1 sugarú és 2^m dimenziójú olyan "Hamming gömb", amelynek a középpontja a 2^{n-m} számú $\mathbf{x} \in C$ kódszó helyén található, és tartalmazza az összes \mathbf{x} -től egy bitben különböző n hosszúságú bináris sorozatokat, összesen 2^n számú n hosszúságú bináris sorozatot fed le, azaz kitölti a teljes $(\mathbb{F}_2)^n$ teret.

Az alábbiakban megmutatjuk, hogy bár a hibajavító dekódolás optimális minimális távolságú (MD) dekódolás a Hamming térben, 3 dB-lel kisebb a kódolási nyeresége, mint a euklideszi térbeli minimális távolságú dekódolásnak (soft decoding).

Tételezzük fel, hogy az (n, k, d) paraméterű bináris lineáris C kód d minimális Hamming távolsága páratlan (a helyzet ekkor rosszabb, mint páros esetben). Legyen \mathbf{x} az átküldött kódszó, és tudjuk, hogy van legalább egy olyan másik kódszó, amelynek a küldött kódszótól mért Hamming távolsága d , és ebből következően van az $s(C)$ -ben legalább egy olyan n sorozat, amelynek az $s(\mathbf{x})$ -től mért négyzetes euklideszi távolsága $4\alpha^2 d$. Bármilyen kis $\varepsilon > 0$ esetén a kemény döntéses dekóder hibázni fog, ha a zaj értéke meghaladja az $\alpha + \varepsilon$ határt a kódszó legalább $(d + 1)/2$ számú olyan helyértékénél, ahol egy másik kódszó különbözik tőle. Ez azt jelenti, hogy kemény döntésnél a döntési küszöbhez tartozó minimális négyzetes távolság az euklideszi térben $\alpha^2(d + 1)/2$ (páros d -nél $\alpha^2 d/2$).

Ezzel szemben "puha" döntésnél (soft decision, or reliability weight) és minimális távolságú (MD) dekódolásnál a dekódolási küszöbtől mért minimális négyzetes távolság az euklideszi térben $\alpha^2 d$. Az uniós korlát pontosságával számolva $Q(\sqrt{\cdot})$ függvény argumentuma $(d + 1)/2d$ -ed részére csökken kemény döntés esetén a puha döntéshez viszonyítva, ami közelítőleg $1/2$ (-3 dB), ha d elegendően nagy (ha d páros, akkor ez az érték pontosan $1/2$).

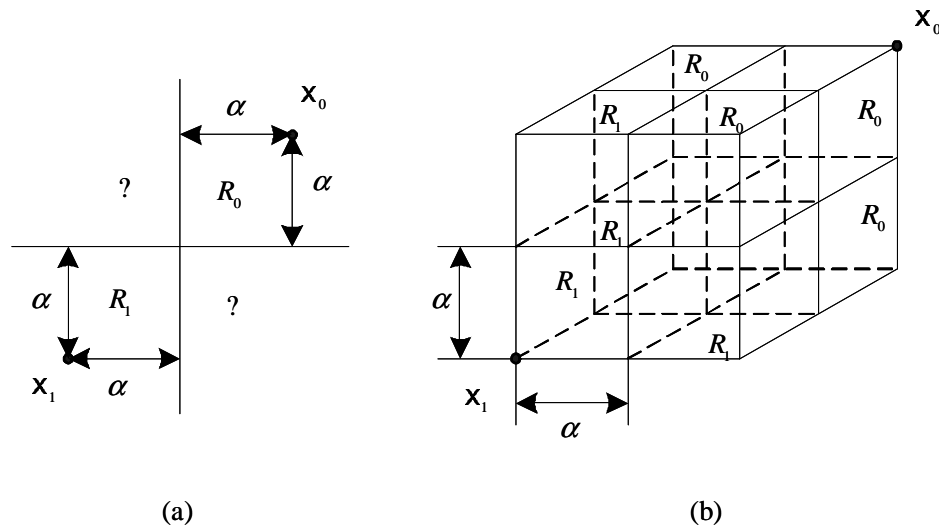
Példa (Az antipodális kódok kemény és puha dekódolása)

Legyen C a $(2, 1, 2)$ bináris kód, ekkor $s(C)$ jeltérben a két kódszónak megfelelő vektorok antipodálisak (lásd A 7.4. (a) ábrát). Kemény döntés esetén, a valós kétdimenziós \mathbb{R}^2 tér négy síknegyedre van osztva, melyeket az egyik vagy másik jelhez (kódszóhoz) rendeljük. Természetesen, azt a két síknegyedet, amelyekben a két jelvektor található, a saját jelvektorokhoz rendeljük, ugyanakkor a másik két síknegyedet nem tudjuk egyértelműen egyik jelhez sem rendelni, ugyanis a jelvektortól α^2 négyzetes távolságra lesz egy döntési határ, amit ha a vett jel zaj hatására átlép, akkor a kódszó egyik komponensében hiba keletkezik. Az ábrán azonban az is látszik, hogy minimális távolságú (MD, puha) döntés esetén a döntési határ négyzetes távolsága mindkét jelvektortól $2\alpha^2$ értékű. Ezért kemény döntés esetén a $P_b = Q(\sqrt{\cdot})$ argumentumában egy 2-es faktor különbség jelenik meg a puha döntés javára, ami pontosan 3 dB veszteséget jelent.

Hasonlóan, ha a C kód a $(3, 1, 3)$ kód, akkor kemény döntés esetén az \mathbb{R}^3 teret 8 térfolyadatra lehet bontani (lásd A 7.4. (b) ábrát). Ebben az esetben (a legegyszerűbb példa erre a Hamming kód) világos, hogy egy-egy jelhez hogyan lehet optimálisan négy-négy térfolyadatot rendelni. A jelvektorokhoz ekkor a legközelebbi döntési határ $2\alpha^2$ távolságra van, ugyanakkor minimális távolságú (MD, puha) döntés esetén a döntési határ négyzetes távolsága mindkét jelvektortól $3\alpha^2$ értékű. Ez most a $P_b \approx Q(\sqrt{\cdot})$ argumentumában egy $2/3$ -as faktor különbség jelenik meg a puha döntés javára, ami pontosan 1,76 dB veszteséget jelent.

Törlés és hibajavítás

Egy olyan dekódolási eljárás, amely a kemény döntés és a "puha" (megbízhatóság alapú) döntés

7.4. ábra. Döntési tartományok az \mathbb{R}^n térben, (a) $(2,1,2)$ kód, (b) $(3,1,3)$ kód

közötti átmenetet jelenti szükségessé teszi a "törlés" fogalmának bevezetését. Ezzel a módszerrel először a vett jelvektor minden r_k összetevőjéhez a $\{0, 1, ?\}$ három értékészletű halmaz elemeit rendeljük az alábbiak szerint:

$$\begin{aligned} r_k &\rightarrow 0 && \text{ha} && r_k > T; \\ r_k &\rightarrow 1 && \text{ha} && r_k < -T; \\ r_k &\rightarrow ? && \text{ha} && -T \leq r_k \leq T, \end{aligned}$$

ahol T a küszöb értéke.

A dekóder ezután lépésről lépésre keresi az így keletkezett három értékészletű n hosszúságú vektorhoz legközelebb eső $\mathbf{x} \in C$ kódszót a Hamming térben, úgy, hogy a kitörölt pozíciókat nem veszi figyelembe a Hamming távolság számításánál. Ha összesen s számú törlés van a vektorban, akkor a kódszavak közötti minimális Hamming távolság $d - s$ lehet a nem törölt pozíciókban, így akkor garantált a korrekt dekódolás, ha a hibák száma t a nem törölt pozíciókban teljesíti a $t < (d - s)/2$, vagy az ekvivalens $2t + s < d$ feltételt. A bináris blokk kódok több osztálya esetén létezik hatékony törléses és hibajavító dekódolási algoritmus, amely garantálja a korrekt dekódolást, ha $2t + s < d$. Az ilyen eljárást korlátozott távolságú törléses és hibajavító dekódolásnak nevezzük.

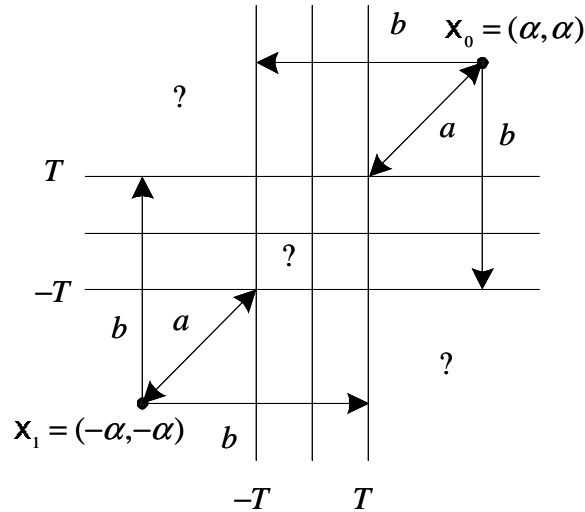
A törléses és hibajavító dekódolást tulajdonképpen a maximális megbízhatóságú (MR) dekódolás egy változatának is tekinthetjük, ahol az összes $|r_k|$ megbízhatósági értéket azonosra állítjuk minden nem törölt helyen, a törölt helyeken pedig 0-ra.

A három szintű kimenet lehetővé teszi, hogy az optimális döntési tartományt jobban meg tudjuk közelíteni az euklideszi térben, mint a kemény döntéssel, és ezzel csökkenteni tudjuk a dekódolási veszteséget. Optimális T küszöbérték esetén a veszteséget tipikusan felére lehet csökkenteni dB-ben.

Példa (Törléses és hibajavító dekódolás)

A 7.5. ábrán bemutatjuk a kemény és/vagy törléses döntés alkalmazásakor keletkező, a $(2,1,2)$ kód minden szimbólumára érvényes 9 döntési tartományt. Három ezek közül bizonytalan. Ezen tartományok minimális négyzetes távolságai a jelvektoroktól a

$$\begin{aligned} a^2 &= 2(\alpha - T)^2; \\ b^2 &= (\alpha + T)^2. \end{aligned}$$



7.5. ábra. Döntési tartományok az \mathbb{R}^n térben a (2,1,2) kód, valamint törlés és kemény döntés esetén

Az a^2 és b^2 minimuma akkor lesz maximális, ha $a^2 = b^2$, amiből az optimális küszöbre a

$$T = \alpha \frac{\sqrt{2} - 1}{\sqrt{2} + 1}$$

értéket kapjuk, amiből

$$a = b = \frac{8}{(\sqrt{2} + 1)^2} \alpha^2 = 1,372 \alpha^2.$$

Ez közelítőleg 1,38 dB-lel jobb, mint a kemény döntéssel elérhető négyzetes euklideszi távolság α^2 , de 1,63 dB-lel rosszabb, mint az MD dekódolással elérhető $2\alpha^2$.

Példa (Törléses és hibajavító dekódolás, optimális T küszöb)

Legyen C egy d minimális távolságú bináris blokk kód, és alkalmazzuk a fent említett törléses és hibajavító dekódolási algoritmust, azaz a vett szimbólumok mindegyikét képezzük le a megismert három értékű halmazra.

- Ekkor bármilyen t és s egész számok esetén, amelyekre érvényes, hogy $2t + s \geq d$ biztosan igaz, hogy a törlések és hibák adott, kedvezőtlen elrendezése esetén dekódolási hiba keletkezhet. Ez az állítás könnyen belátható, ha feltételezzük, hogy éppen az $\mathbf{x} \in C$ kódszót küldték, és van egy $\mathbf{y} \in C$ kódszó, melynek az \mathbf{x} -től mért $d_H(\mathbf{x}, \mathbf{y})$ Hamming távolsága azonos d -vel.

Tételezzük fel, hogy a dekódolás során az s számú törlés azokon a helyeken történt, amelyeken az \mathbf{x} és \mathbf{y} kódszavak különböznek egymástól. A törlések után a megmaradó kódszavak távolsága ekkor éppen $d - s$ lesz. Ha a feltételek szerint a hibák száma $t \geq (d - s)/2$, és ezek a hibák a kemény döntéseknél éppen azokon a helyeken keletkeztek, ahol az \mathbf{x} és \mathbf{y} kódszavak különböznek egymástól akkor biztos, hogy a törléses és hibajavító dekódolási algoritmus során keletkezett bináris vektor közelebb lesz az \mathbf{y} , mint az \mathbf{x} kódszóhoz, vagy a két távolság azonos lesz, ami azt jelenti, hogy a dekódolás során hiba keletkezhet;

- A minimális négyzetes euklideszi távolság bármely $\mathbf{x} \in C$ kódszótól a $\min_{2t+s \geq d} = \min(s(\alpha - T)^2 + t(\alpha + T)^2)$ kifejezéssel határozható meg.

Ez abból következik, hogy az $s(C)$ konstellációban a jelvektor s különböző komponensénél a törléshez tartozó döntési határ $(\alpha - T)^2$ négyzetes távolságra van az $s(\mathbf{x})$ jelvektortól, ami azt jelenti, hogy s számú törléshez legalább az $s(\alpha - T)^2$ négyzetes küszöböt kell a zajnak meghaladnia.

Ugyanakkor az $s(C)$ konstellációban a jelvektor t különböző komponensénél a hibás kemény döntéshez tartozó döntési határ $(\alpha + T)^2$ négyzetes távolságra van az $s(\mathbf{x})$ jelvektortól, ami azt jelenti, hogy t számú hibához legalább az $t(\alpha + T)^2$ négyzetes küszöböt kell a zajnak meghaladnia.

A két négyzetes küszöb összege akkor lesz a $2t + s \geq d$ feltétel mellett minimális, ha $2t + s = d$, mivel például bármely rögzített t esetén az $s = d - 2t$ választás mellett kapjuk a minimális értéket (ha $\alpha > T$). Éppen ezért a minimum kereséséhez deriválni kell a $(d - 2t)(\alpha - T)^2 + t(\alpha + T)^2$ kifejezést t szerint, ami alapján a minimum helyét az

$$2(\alpha - T)^2 = (\alpha + T)^2$$

egyenlőségéből határozhatjuk meg. Ebből az optimális küszöb értékére általánosan is a

$$T = \alpha \frac{\sqrt{2} - 1}{\sqrt{2} + 1}$$

eredmény adódik.

- Ennél a T értéknél a minimális euklideszi távolság

$$d_{\min} = (d - 2t)(\alpha - T)^2 + t(\alpha + T)^2 = d\alpha^2 \left(1 - \frac{\sqrt{2} - 1}{\sqrt{2} + 1}\right)^2 = d\alpha^2 \frac{4}{(\sqrt{2} + 1)^2} = 0,686\alpha^2 d.$$

Az MD dekódoláshoz képesti veszteség most is 1,63 dB.

Általánosított minimális távolságú dekódolás

Az eddigi lépésekkel egyre közelebb kerültünk az optimális MD dekódoláshoz az uniós korlát pontosságát figyelembe véve. A következő lépés az úgynevezett általánosított minimális távolságú (generalized minimum distance decoding, GMD) dekódolás.

GMD dekódolás esetén a dekóder elvégzi a kemény és puha döntést is, azaz bináris szimbólumként egyszerre állítja elő a vett szimbólum $\text{sgn}(r_k)$ előjelét és az $|r_k|$ megbízhatóság értéket, és a kemény döntések eredményeit a megbízhatóság alapján sorba rendezi.

A GMD dekóder ezután végrehajt egy törléses és hibajavító dekódolási sorozatot, mely során rendre törli az $s = d - 1, d - 3, \dots$ legkisebb megbízhatóságú szimbólumokat. (A többi értékek törlése nem szükséges, mivel ha $d - s$ páros és $2t < d - s$, akkor az $2t < d - s - 1$ reláció is fennáll, ezért egy ilyen kísérlettel eggyel több szimbólumot fogunk törölni, de a dekódoló ugyanazt a kódszót fogja választani).

A lépések száma $d/2$, ha d páros és $(d + 1)/2$, ha d páratlan, vagyis a dekódolás során $d/2$ lépésre van szükség. Minden lépés során az eljárás előállít egy esélyes kódszót. Ezért $d/2$ lépés esetén éppen $d/2$ számú különböző kódszó keletkezhet. Ezeket a kódszavakat ezután össze lehet hasonlítani például az $r(\mathbf{x}|\mathbf{r})$ megbízhatóságuk alapján, és ki lehet választani közülük a legjobbat.

Példa (Általánosított minimális távolságú (GMD) dekódolás)

Egy $(n, n - 1, 2)$ paraméterű SPC kód esetén a GMD dekódolás során csak egyetlen kísérletet kell végrehajtani, amikor a legkisebb megbízhatóságú szimbólumot töröljük. Az így keletkező esélyes kódszó a nem törölt pozíciók kívül biztosan megegyezik egy valódi kódszóval. A GMD dekódolás ebben az esetben azonos a Wagner dekódolási szabállyal, amiről tudjuk, hogy optimális.

Belátható, hogy a GMD dekódolás során nem keletkezik hiba, ha a zaj négyzetes normája $\|\mathbf{n}\|^2$ nem haladja meg az $\alpha^2 d$ értéket, ugyanis a döntési határ és a vektortér elemeinek a négyzetes távolsága

éppen $\alpha^2 d$, mint az MD dekódolásnál. Ilyenkor nincsen különbség az MD és a GMD dekódolás hibaexponense között.

A vizsgálatok általában kimutatták, hogy az algebrai blokk kódok fontosabb osztályainál a GMD dekódolás komplexitása valamivel nagyobb, mint a kemény döntésé vagy a törléses és hibajavító dekódolásé. Továbbá azt is igazolták, hogy nemcsak a hibaexponens, hanem az uniós korlát hibakoefficiense K_d is azonos az MD és a GMD dekódolás esetén.

A GMD dekódolás korlátozott távolságú dekódolás, mivel a döntési tartományok $\alpha^2 d$ sugarú gömbök, amelyek a MD dekódolás R_j döntési tartományán belül helyezkednek el. Így a GMD dekódolás teljesítőképessége az MD dekódolás alatt marad, de jobb a törléses és hibajavító dekódolásnál ≤ 1 dB-lel. Ez az oka annak, hogy a GMD dekódolást a gyakorlatban ritkán alkalmazzák.

8. fejezet

A modulált jelek spektrális vizsgálata

A modulált jeleket általában olyan csatornán vesszük át, amelyben a sávszélesség csak korlátozottan áll a rendelkezésünkre. Éppen ezért fontos ismerni a modulált jelek teljesítménysűrűség-függvényét, vagyis azt, hogy a frekvenciasáv egyes részterületeiben mekkora a jel egységnyi sávszélességre jutó teljesítménye. Fontos ez azért, mert

- csak ennek ismeretében tudjuk meghatározni, hogy a különböző forrásokból származó jelek mekkora interferenciát okoznak egy másik csatornában;
- csak így tudjuk illeszteni a jelet a rendelkezésre álló frekvenciasávhoz.

8.1. A ciklostacionárius jelek tulajdonságai

Egy sztochasztikus folyamathoz akkor tudunk teljesítménysűrűség-függvényt rendelni, ha a folyamat legalább gyengén stacioner, azaz

$$\mathbf{E}\{x(t)\} = \text{konst.} \quad \text{és} \quad \mathbf{E}\{x(t)x(t+\tau)\} = R(\tau),$$

vagyis a folyamat első és második momentumai nem függenek az időtől.

Ebben az esetben a teljesítménysűrűség-függvényt az

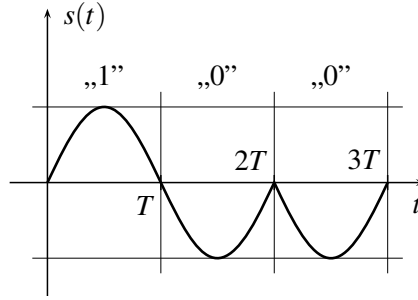
$$s(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R(\tau) e^{-j\omega\tau} d\tau \quad \left[\frac{\text{W}}{\frac{\text{rad}}{\text{sec}}} \right],$$
$$s(f) = \int_{-\infty}^{\infty} R(\tau) e^{-j2\pi f\tau} d\tau \quad \left[\frac{\text{W}}{\text{Hz}} \right]$$

kifejezésekkel határozhatjuk meg. Természetesen

$$R(\tau) = \int_{-\infty}^{\infty} s(\omega) e^{j\omega\tau} d\omega,$$
$$R(\tau) = \int_{-\infty}^{\infty} s(f) e^{j2\pi f\tau} df.$$

A modulált jelek mint sztochasztikus folyamatok általában nem teljesítik a gyenge stacionaritás feltételeit, mivel időbeli ütemezésükben periodicitások vannak (szimbólumidő, periodikus vivő, stb.). Ennek magyarázatához tekintsük a következő PAM típusú modulált jelet:

$$s(t) = \sum_{i=-\infty}^{\infty} \xi_{ki} g_T(t - iT),$$



8.1. ábra. A PAM jel egy realizációja

és legyen $\xi_{k_i} \in \{-1, 1\}$, $\pi_1 = \pi_{-1} = \frac{1}{2}$, valamint

$$g_T(t) = \begin{cases} \sin\left(\frac{\pi t}{T}\right), & t \in [0, T) \\ 0 & \text{egyébként} \end{cases}$$

A jel realizációját a 8.1. ábra mutatja.

Az ábrából nyilvánvaló, hogy ennek a jelnek a momentumai függenek az időtől, hiszen

$$\mathbf{E}\{s(0)s(0+\tau)\} = 0, \quad \forall \tau,$$

de

$$\mathbf{E}\left\{s\left(\frac{T}{2}\right)s\left(\frac{T}{2}+\tau\right)\right\} = \frac{1}{2}\left(\sin\left(\frac{\pi}{2}+\frac{\pi\tau}{T}\right)+(-1)\left(-\sin\left(\frac{\pi}{2}+\frac{\pi\tau}{T}\right)\right)\right), \quad |\tau| < \frac{T}{2}.$$

Ebből az következik, hogy ennek a jelnek nem értelmezhető a teljesítménysűrűség-függvénye.

Az is világosan látható viszont, hogy a jel ciklikusan stacioner, azaz fennállnak az alábbi tulajdonságai:

$$\mathbf{E}\{s(t)\} = \mathbf{E}\{s(t+kT)\}, \quad k \in \mathbb{Z},$$

$$\mathbf{E}\{s(t)s(t+\tau)\} = \mathbf{E}\{s(t+kT)s(t+kT+\tau)\}, \quad k \in \mathbb{Z},$$

ami azt jelenti, hogy a jel első és második momentumai az idő periodikus függvényei. Az ilyen jeleket ciklostacioner jeleknek nevezzük. Elmondható, hogy a rögzített időraszterrel rendelkező jelek (szimbólumidőzítés, rögzített vivőfázis) ciklostacioner jelek, ezért a teljesítménysűrűség spektrumuk nem értelmezhető.

A probléma megoldása az, hogy a modulált jeleket, általában a ciklostacioner sztochasztikus folyamatok realizációhalmazát kibővítjük az eredeti realizációknak egy perióduson belüli időben véletlenül eltolt változatával, vagyis példánkban az

$$s(t) = \sum_{i=-\infty}^{\infty} \xi_{k_i} g_T(t-iT)$$

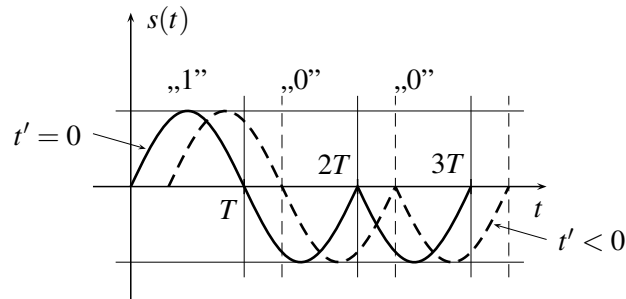
helyett az

$$s(t) = \sum_{i=-\infty}^{\infty} \xi_{k_i} g_T(t+t'-iT)$$

jelet tekintjük (ahol t' egy $\{\xi_{k_i}\}$ -ktől független, a $[0, T)$ tartományban egyenletes eloszlású valószínűségi változó), melynek realizációit a 8.2. ábra illusztrálja.

Igen egyszerűen belátható, hogy az így generált sztochasztikus folyamat már stacioner, ha a $\{\xi_{k_i}\}$ sorozat stacioner. Ily módon lehetőség van arra, hogy a folyamat teljesítménysűrűség-függvényét meghatározzuk. Ehhez először ki kell számítani a jel

$$\begin{aligned} R(\tau) &= \mathbf{E}_{t'} \left\{ \mathbf{E}_{\xi_{k_i}} \{s(t)s(t+\tau)\} \right\} = \\ &= \frac{1}{T} \int_0^T \mathbf{E}_{\xi_{k_i}} \{s(t)s(t+\tau)\} dt' \end{aligned}$$



8.2. ábra. Ciklostacioner jelek időben véletlenül eltolt realizációi

korrelációs függvényét, majd ebből kalkulálható az $s(\omega)$.

8.2. A véletlen fázisú szinuszos jel teljesítménysűrűség-függvénye

Az előbbieket illusztrálására határozzuk meg egy véletlen fázisú vivőfrekvenciás jel teljesítménysűrűség-függvényét. Az

$$s(t) = \sqrt{2} \cos(\omega_0 t), \quad P = 1$$

periodikus jel nem stacioner (sőt, nem is sztochasztikus folyamat), de az

$$s(t) = \sqrt{2} \cos(\omega_0 t + \varphi), \quad P = 1,$$

ahol φ egy $[0, 2\pi)$ -ben egyenletes eloszlású valószínűségi változó, már gyengén stacioner, hiszen

$$\mathbf{E}_\varphi \{s(t)\} = \frac{1}{2\pi} \int_0^{2\pi} \sqrt{2} \cos(\omega_0 t + \varphi) d\varphi = 0$$

$$\begin{aligned} \mathbf{E}_\varphi \{s(t)s(t+\tau)\} &= \frac{1}{2\pi} \int_0^{2\pi} 2 \cos(\omega_0 t + \varphi) \cos(\omega_0(t+\tau) + \varphi) d\varphi = \\ &= \frac{1}{2\pi} \int_0^{2\pi} 2 \left(\frac{\cos(2\omega_0 t + \omega_0 \tau + 2\varphi)}{2} + \frac{\cos(\omega_0 \tau)}{2} \right) d\varphi = \\ &= \cos(\omega_0 \tau) \end{aligned}$$

nem függvénye az időnek, továbbá tudjuk, hogy

$$R(\tau) = \cos(\omega_0 \tau) = \frac{e^{j\omega_0 \tau} + e^{-j\omega_0 \tau}}{2},$$

és ebből

$$\begin{aligned} s(\omega) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} R(\tau) e^{-j\omega \tau} d\tau = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{j\omega_0 \tau} + e^{-j\omega_0 \tau}}{2} e^{-j\omega \tau} d\tau = \\ &= \frac{1}{2} \delta(\omega + \omega_0) + \frac{1}{2} \delta(\omega - \omega_0). \end{aligned}$$

Tudjuk, hogy a jel teljes teljesítményét a

$$P = \int_{-\infty}^{\infty} s(\omega) d\omega = R(0) = 1$$

adja, ami a kiinduló szinuszos jelre természetesen teljesül.

8.3. Az alapsávi PAM jelek teljesítménysűrűség-függvénye

Vizsgáljuk meg az

$$s(t) = \sum_{i=-\infty}^{\infty} \xi_{k_i} g_T(t + t' - iT)$$

jel spektrumát.

$$\begin{aligned} R(\tau) &= \mathbf{E}_{t'} \left\{ \mathbf{E}_{\xi_{k_i}} \{s(t)s(t+\tau)\} \right\} = \\ &= \frac{1}{T} \int_0^T \mathbf{E}_{\xi_{k_i}} \{s(t)s(t+\tau)\} dt' = \\ &= \frac{1}{T} \int_0^T \mathbf{E} \left\{ \left(\sum_{i=-\infty}^{\infty} \xi_{k_i} g_T(t+t'-iT) \right) \left(\sum_{j=-\infty}^{\infty} \xi_{k_j} g_T(t+t'+\tau-jT) \right) \right\} dt' = \\ &= \frac{1}{T} \int_0^T \mathbf{E} \left\{ \left(\sum_{i=-\infty}^{\infty} \xi_{k_i} g_T(t-iT) \right) \left(\sum_{j=-\infty}^{\infty} \xi_{k_j} g_T(t+\tau-jT) \right) \right\} dt = \\ &= \frac{1}{T} \int_0^T \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \mathbf{E} \{ \xi_{k_i} \xi_{k_j} \} g_T(t-iT) g_T(t+\tau-jT) dt \end{aligned}$$

Tételezzük fel, hogy a $\{\xi_{k_i}\}$ sorozat stacioner, független és nulla várható értékű, azaz

$$\mathbf{E} \{ \xi_{k_i} \} = \mathbf{E} \{ \xi \} = 0$$

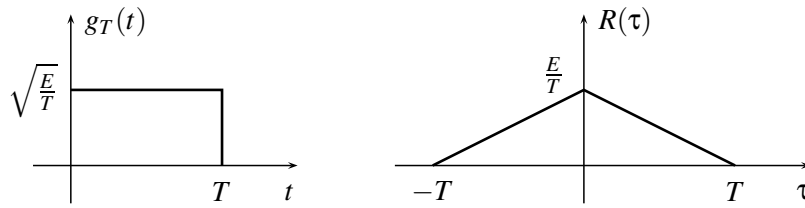
$$\mathbf{E} \{ \xi_{k_i} \xi_{k_j} \} = \begin{cases} \mathbf{E} \{ \xi^2 \}, & \text{ha } i = j \\ 0 & \text{egyébként} \end{cases}$$

akkor

$$\begin{aligned} R(\tau) &= \frac{\mathbf{E} \{ \xi^2 \}}{T} \int_0^T \sum_{i=-\infty}^{\infty} g_T(t-iT) g_T(t+\tau-iT) dt = \\ &= \frac{\mathbf{E} \{ \xi^2 \}}{T} \sum_{i=-\infty}^{\infty} \int_{-iT}^{T-iT} g_T(t) g_T(t+\tau) dt = \\ &= \frac{\mathbf{E} \{ \xi^2 \}}{T} \int_{-\infty}^{\infty} g_T(t) g_T(t+\tau) dt. \end{aligned}$$

Ebből a jel teljesítménysűrűség-függvénye az alábbi módon számolható:

$$s(\omega) = \frac{\mathbf{E} \{ \xi^2 \}}{T} \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} g_T(t) g_T(t+\tau) dt \right) e^{-j\omega\tau} d\tau =$$



8.3. ábra. Az elemi jelalak és az autokorrelációs függvény NRZ moduláció esetén

$$\begin{aligned}
 &= \frac{\mathbf{E}\{\xi^2\}}{T} \frac{1}{2\pi} \int_{-\infty}^{\infty} g_T(t) \left(\int_{-\infty}^{\infty} g_T(t+\tau) e^{-j\omega\tau} d\tau \right) dt = \\
 &= \frac{\mathbf{E}\{\xi^2\}}{T} \frac{1}{2\pi} \int_{-\infty}^{\infty} g_T(t) G_T(\omega) e^{j\omega t} dt = \\
 &= \frac{\mathbf{E}\{\xi^2\}}{2\pi T} G_T(\omega) G_T^*(\omega) = \\
 &= \frac{\mathbf{E}\{\xi^2\}}{2\pi T} |G_T(\omega)|^2,
 \end{aligned}$$

ahol

$$G_T(\omega) = \int_{-\infty}^{\infty} g_T(t) e^{-j\omega t} dt$$

az elemi adójel Fourier-transzformáltja.

8.4. Illusztratív példák a PAM jelek spektrális analízisére

Az NRZ moduláció spektruma

$$\xi \in \{-1, 1\},$$

NRZ jelválasztás esetén (8.3. ábra)

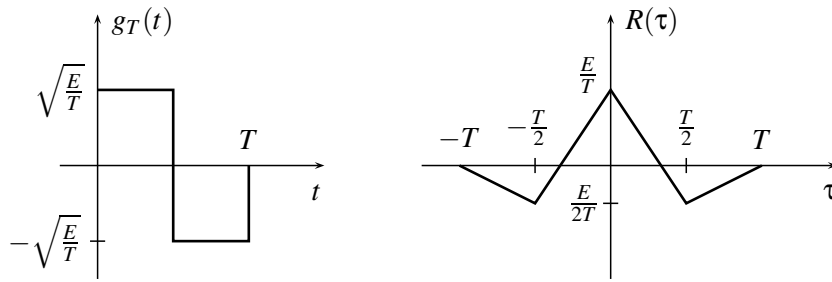
$$g_T(t) = \begin{cases} \sqrt{\frac{E}{T}}, & \text{ha } t \in [0, T) \\ 0 & \text{egyébként} \end{cases}$$

Ekkor

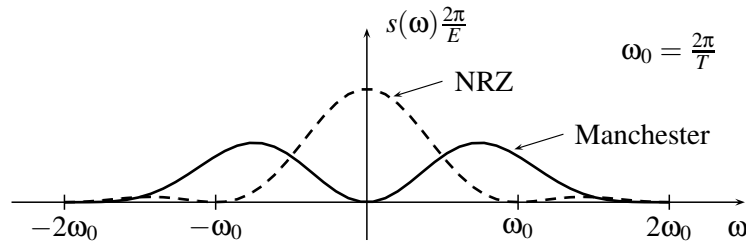
$$R(\tau) = \begin{cases} \frac{E}{T} \left(1 - \frac{|\tau|}{T}\right), & \text{ha } |\tau| \in [0, T) \\ 0 & \text{egyébként} \end{cases}$$

és

$$s(\omega) = \frac{E}{2\pi} \left(\frac{\sin\left(\frac{\omega T}{2}\right)}{\frac{\omega T}{2}} \right)^2.$$



8.4. ábra. Az elemi jelalak és az autokorrelációs függvény Manchester-kódolás esetén



8.5. ábra. Az NRZ és a Manchester-kódolás spektruma

A Manchester-kódot alkalmazó jel spektruma

$$\xi \in \{-1, 1\},$$

Manchester-kód esetén (8.4. ábra)

$$g_T(t) = \begin{cases} \sqrt{\frac{E}{T}}, & \text{ha } t \in [0, \frac{T}{2}) \\ -\sqrt{\frac{E}{T}}, & \text{ha } t \in [\frac{T}{2}, T) \\ 0 & \text{egyébként} \end{cases}$$

Ekkor

$$R(\tau) = \begin{cases} \frac{E}{T} \left(1 - 3\frac{|\tau|}{T}\right), & \text{ha } |\tau| \in [0, \frac{T}{2}) \\ \frac{E}{T} \left(-1 + \frac{|\tau|}{T}\right), & \text{ha } |\tau| \in [\frac{T}{2}, T) \\ 0 & \text{egyébként} \end{cases}$$

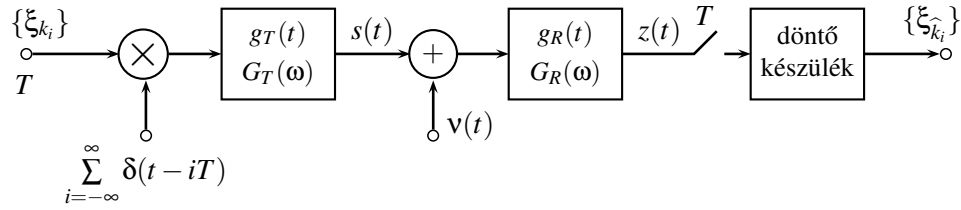
és

$$s(\omega) = \frac{E}{2\pi} \frac{\sin^4\left(\frac{\omega T}{4}\right)}{\left(\frac{\omega T}{4}\right)^2}.$$

A két spektrumot tartalmazó 8.5. ábrából megállapítható, hogy az NRZ kódolt jel lényeges energiájú komponenseket tartalmaz a nulla körüli frekvenciatartományban, ezért átvitelére nem alkalmas olyan csatorna, amelyben nincs DC csatolás (pl. telefoncsatorna). A Manchester-kódolt jel átvihető felüláteresztő (DC átvitelre nem alkalmas) csatornán is. Mivel itt az elemi jelnek nulla az átlagértéke, így a modulált jelben sincs DC összetevő.

8.5. Az általános optimális PAM rendszer vizsgálata szimbólumközi áthallásmentes esetben

Az optimális PAM típusú rendszer felépítését a 8.6. ábra tartalmazza. Tételezzük fel, hogy a vevőszűrő kimenetén lévő jelre teljesül a Nyquist-feltétel, azaz a döntőkészülék egy mintából éppen egy



8.6. ábra. Az optimális PAM típusú rendszer felépítése

szimbólumot tud becsülni.

A korábbiakból tudjuk azt is, hogy az optimális vevőben illesztett szűrőt (korrelációs detektort) kell megvalósítani, azaz

$$g_R(t) = g_T(-t),$$

$$\begin{aligned} G_R(\omega) &= \int_{-\infty}^{\infty} g_R(t) e^{-j\omega t} dt = \\ &= \int_{-\infty}^{\infty} g_T(-t) e^{-j\omega t} dt = \\ &= \int_{-\infty}^{\infty} g_T(t) e^{j\omega t} dt = \\ &= G_T^*(\omega), \end{aligned}$$

így az

$$s(t) = \sum_{k_i} \xi_{k_i} g_T(t - iT)$$

és

$$z(t) = \sum_{k_i} \xi_{k_i} g(t - iT),$$

ahol

$$\begin{aligned} g(t) &= g_T(t) * g_R(t), \\ G(\omega) &= G_T(\omega) G_R(\omega) = |G_T(\omega)|^2. \end{aligned}$$

Természetesen a szimbólumközi áthallásmentességnek a $z(t)$ jelre kell teljesülnie, ami a $G(\omega)$ Nyquist-ekvivalensére ad megkötést, miszerint legyen

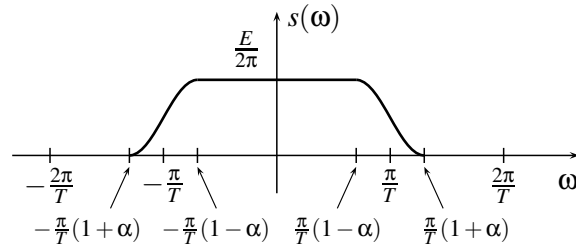
$$G_{\text{ekv}}(\omega) = \begin{cases} c, & \text{ha } |\omega| \leq \frac{\pi}{T} \\ 0 & \text{egyébként} \end{cases}$$

(c helyett lehet $c e^{j\omega T_0}$ is), ahol

$$G_{\text{ekv}}(\omega) = \begin{cases} \sum_{j=-\infty}^{\infty} G(\omega - j\omega_0), & \text{ha } |\omega| \leq \frac{\pi}{T} \\ 0 & \text{egyébként} \end{cases}$$

Az eddigiekből megállapítható, hogy optimális, illesztett szűrős vevőben, ahol teljesül a szimbólumközi áthallásmentesség, $G(\omega) = |G_T(\omega)|^2$ teljesíti a Nyquist-feltételt.

$$s(\omega) = \frac{\mathbf{E}\{\xi^2\}}{2\pi T} |G_T(\omega)|^2$$



8.7. ábra. Az emelt koszinuszos jel spektruma

az adó kimenetén lévő jel teljesítménysűrűség-függvénye, azaz az adó kimenetén lévő jel spektruma a vevőszűrő kimenetén megjelenő $g(t)$ elemi jel Fourier-transzformáltjának abszolút értékével arányos.

Így például abban az esetben, ha $G(\omega)$ ún. emelt koszinusz függvény α lekerekítési paraméterrel, akkor

$$G(\omega) = \begin{cases} ET, & \text{ha } |\omega| \leq \frac{\pi}{T}(1-\alpha) \\ ET \frac{1}{2} \left(1 - \sin\left(\frac{T}{2\alpha}\left(|\omega| - \frac{\pi}{T}\right)\right)\right), & \text{ha } \frac{\pi}{T}(1-\alpha) < |\omega| \leq \frac{\pi}{T}(1+\alpha) \\ 0 & \text{egyébként} \end{cases}$$

$G_T(\omega) = \sqrt{G(\omega)}$ és $\xi \in \{-1, 1\}$ esetén

$$s(\omega) = \frac{1}{2\pi T} |G_T(\omega)|^2 = \frac{1}{2\pi T} G(\omega)$$

(lásd a 8.7. ábrán). Megjegyezzük, hogy esetünkben teljesülnie kell annak, hogy

$$R(0) = \int_{-\infty}^{\infty} s(\omega) d\omega = \frac{1}{2\pi T} \int_{-\infty}^{\infty} |G_T(\omega)|^2 d\omega = \frac{E}{T}$$

a jel teljesítménye.

8.6. Részleges válaszfüggvényű PAM típusú rendszerek

A modulált jelek teljesítménysűrűség-függvénye alapvetően megváltoztatható, ha a rendszerben megengedünk — jól kézbentartott módon — szimbólumközi áthallást, azaz azt, hogy a vevőszűrő kimenetén lévő jel mintái egynél több átküldött szimbólumtól függjenek. Ilyenkor a minták értékét több szimbólum lineáris kombinációja határozza meg. Legyen a rendszerünkben a vevőszűrő kimenetén lévő jel

$$z(t) = \sum_{i=-\infty}^{\infty} \xi_{ki} g(t - iT),$$

ahol

$$g(t) := a_0 g_0(t) + a_1 g_0(t - T) + \dots + a_P g_0(t - PT) = \sum_{l=0}^P a_l g_0(t - lT).$$

Válasszuk $g_0(t)$ -t úgy, hogy teljesítse a Nyquist-feltételt, azaz

$$g_0(kT) = \begin{cases} g_0(0), & k = 0 \\ 0, & k \neq 0 \end{cases}$$

Ekkor

$$z(t) = \sum_{i=-\infty}^{\infty} \xi_{ki} g(t - iT) = \sum_{i=-\infty}^{\infty} \xi_{ki} \sum_{l=0}^P a_l g_0(t - lT - iT),$$

és vizsgáljuk meg $z(t)$ mintáit.

$$\begin{aligned}
 z(nT) &= \sum_{i=-\infty}^{\infty} \xi_{k_i} g(nT - iT) = \\
 &= \sum_{i=-\infty}^{\infty} \xi_{k_i} \sum_{l=0}^P a_l g_0(nT - lT - iT) = \\
 &= g_0(0) (a_0 \xi_{k_n} + a_1 \xi_{k_{n-1}} + \dots + a_P \xi_{k_{n-P}}) = \\
 &= g_0(0) \sum_{l=0}^P a_l \xi_{k_{n-l}}
 \end{aligned}$$

Megállapíthatjuk, hogy az n -edik mintára az aktuális n -edik szimbólumon kívül még P számú korábbi szimbólum is hat, ami a döntést nyilvánvalóan megnehezíti. Az ilyen rendszerekben a vevő komplexitása megnő.

Esetünkben a vevőszűrő kimenetén lévő elemi jelre igaz, hogy

$$G(\omega) = \sum_{l=0}^P a_l \mathcal{F}(g_0(t - lT)) = \sum_{l=0}^P a_l G_0(\omega) e^{-j\omega lT},$$

amiből

$$s(\omega) = \frac{\mathbf{E}\{\xi^2\}}{2\pi T} |G(\omega)|.$$

8.7. Példák a részleges válaszfüggvényű rendszerek spektrális vizsgálata

Duobináris jel

$$P = 1, a_0 = 1, a_1 = 1, \xi \in \{-1, 1\}$$

$$G_0(\omega) = \begin{cases} c, & |\omega| \leq \frac{\pi}{T} \\ 0, & |\omega| > \frac{\pi}{T} \end{cases}$$

$$G(\omega) = \begin{cases} c(1 + e^{-j\omega T}) = 2ce^{-j\omega \frac{T}{2}} \cos(\omega \frac{T}{2}), & |\omega| \leq \frac{\pi}{T} \\ 0, & |\omega| > \frac{\pi}{T} \end{cases}$$

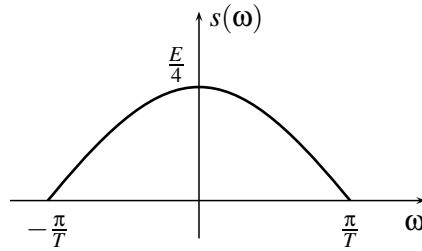
Megjegyezzük, hogy az $e^{-j\omega \frac{T}{2}}$ faktor egyszerűen elhagyható, hiszen ez csak egy $\frac{T}{2}$ idejű késleltést jelent, vagyis

$$G(\omega) = \begin{cases} 2c \cos(\omega \frac{T}{2}), & |\omega| \leq \frac{\pi}{T} \\ 0, & |\omega| > \frac{\pi}{T} \end{cases}$$

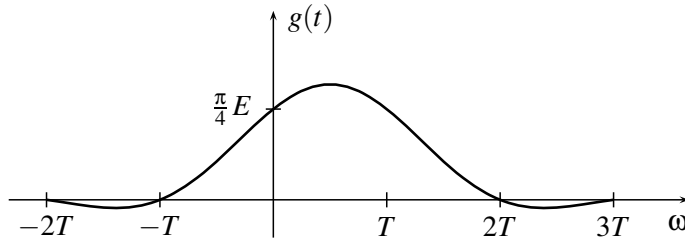
így

$$G_T(\omega) = \begin{cases} \sqrt{2c \cos(\omega \frac{T}{2})}, & |\omega| \leq \frac{\pi}{T} \\ 0, & |\omega| > \frac{\pi}{T} \end{cases}$$

$$s(\omega) = \frac{1}{2\pi T} \begin{cases} 2c \cos(\omega \frac{T}{2}), & |\omega| \leq \frac{\pi}{T} \\ 0, & |\omega| > \frac{\pi}{T} \end{cases}$$



8.8. ábra. Részleges válaszfüggvényű rendszer teljesítménysűrűség-függvénye $a_0 = 1, a_1 = 1$ paraméterű duobináris jel esetén



8.9. ábra. A vevőszűrő kimenetén megjelenő jel $a_0 = 1, a_1 = 1$ paraméterű duobináris jel esetén

Mivel

$$\int_{-\infty}^{\infty} s(\omega) d\omega = \frac{E}{T} = \int_{-\frac{\pi}{T}}^{\frac{\pi}{T}} \frac{c}{\pi T} \cos\left(\omega \frac{T}{2}\right) d\omega = \frac{4c}{\pi T^2},$$

ezért $c = \frac{\pi}{4} ET$. A rendszer teljesítménysűrűség-függvénye a 8.8. ábrán, míg a vevőszűrő kimenetén megjelenő jel a 8.9. ábrán látható.

Duobináris jel

$$P = 1, a_0 = 1, a_1 = -1, \xi \in \{-1, 1\}$$

$$G_0(\omega) = \begin{cases} c, & |\omega| \leq \frac{\pi}{T} \\ 0, & |\omega| > \frac{\pi}{T} \end{cases}$$

$$G(\omega) = \begin{cases} c(1 - e^{-j\omega T}) = 2jce^{-j\omega \frac{T}{2}} \sin\left(\omega \frac{T}{2}\right), & |\omega| \leq \frac{\pi}{T} \\ 0, & |\omega| > \frac{\pi}{T} \end{cases}$$

Ebből

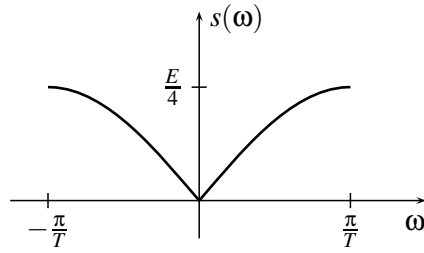
$$|G(\omega)| = \begin{cases} 2c \sin\left(|\omega| \frac{T}{2}\right), & |\omega| \leq \frac{\pi}{T} \\ 0, & |\omega| > \frac{\pi}{T} \end{cases}$$

így

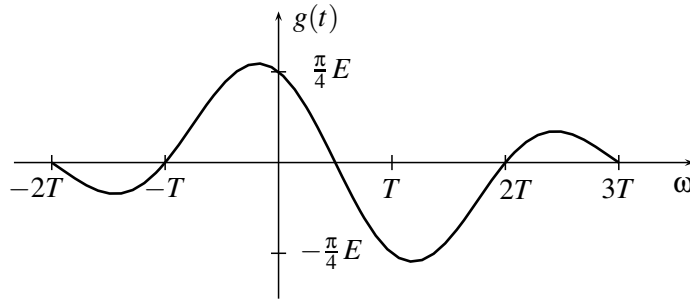
$$s(\omega) = \frac{1}{2\pi T} \begin{cases} 2c \sin\left(|\omega| \frac{T}{2}\right), & |\omega| \leq \frac{\pi}{T} \\ 0, & |\omega| > \frac{\pi}{T} \end{cases}$$

és $c = \frac{\pi}{4} ET$. A rendszer teljesítménysűrűség-függvénye a 8.10. ábrán, míg a vevőszűrő kimenetén megjelenő jel a 8.11. ábrán látható.

A megoldás előnye, hogy a jel átvihető DC csatolás nélküli csatornán (pl. telefoncsatornán), mivel $g(t)$ átlagértéke 0. A jel sávzélessége a $[-\frac{\pi}{T}, \frac{\pi}{T}]$ tartományban korlátozott, tehát keskenyebb, mint az azonos tulajdonságú Manchester-kódolt jelé.



8.10. ábra. Részleges válaszfüggvényű rendszer teljesítménysűrűség-függvénye $a_0 = 1, a_1 = -1$ paraméterű duobináris jel esetén



8.11. ábra. A vevőszűrő kimenetén megjelenő jel $a_0 = 1, a_1 = -1$ paraméterű duobináris jel esetén

Duobináris jel véges időbeli tartójú elemi jellel

$$P = 1, a_0 = 1, a_1 = 1, \xi \in \{-1, 1\}$$

$$g_T(t) = \begin{cases} \sqrt{\frac{1}{2} \frac{E}{T}}, & t \in [0, 2T) \\ 0 & \text{egyébként} \end{cases}$$

$$G_T(\omega) = \sqrt{\frac{E}{2T}} 2T e^{-j\omega T} \frac{\sin(\omega T)}{\omega T}$$

$$|G_T(\omega)|^2 = 2ET \frac{\sin^2(\omega T)}{(\omega T)^2}$$

$$s(\omega) = \frac{1}{2\pi T} 2ET \frac{\sin^2(\omega T)}{(\omega T)^2} = \frac{E}{\pi} \left(\frac{\sin(\omega T)}{\omega T} \right)^2$$

A jel sávszélessége keskenyebb, mint az NRZ kódolt jelé.

8.8. Általános modulációs rendszer vizsgálata tetszőleges elemi jelekkel és független szimbólumsorozattal

Feltételek:

A forrásból érkező szimbólumsorozat $\{\xi_{k_i}\}$, $k_i \in \{1, \dots, M\}$, $i \in \mathbb{Z}$, π_k a priori valószínűségekkel. A k_i -edik szimbólumhoz az i -edik időrészben az $x_{k_i}(t - iT)$ jelet rendeljük, ezért a modulált jel

$$s(t) = \sum_{i=-\infty}^{\infty} x_{k_i}(t - iT).$$

Az $\{x_k(t)\}$ elemi jelek Fourier-transzformáltját az

$$X_k(\omega) = \int_{-\infty}^{\infty} x_k(t) e^{-j\omega t} dt$$

$$x_k(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X_k(\omega) e^{j\omega t} d\omega$$

kifejezésekkel számolhatjuk. Az $s(t)$ korrelációs függvényét pedig az

$$R(\tau) = \frac{1}{T} \int_0^T \mathbf{E} \{s(t)s(t+\tau)\} dt$$

kifejezés adja meg.

A korrelációs függvény számítása előtt bontsuk fel az $s(t)$ jelet két részre az alábbiak szerint:

$$s(t) = \sum_{i=-\infty}^{\infty} x_{k_i}(t-iT) = \sum_{i=-\infty}^{\infty} (x'_{k_i}(t-iT) + \bar{x}(t-iT)),$$

ahol

$$\bar{x}(t) = \mathbf{E} \{x_k(t)\} = \sum_{k=1}^M \pi_k x_k(t)$$

az átlagos elemi jel és

$$x'_k(t) = x_k(t) - \bar{x}(t)$$

az elemi jel és az átlagos elemi jel különbsége.

Nyilvánvaló, hogy az

$$s_2(t) = \sum_{i=-\infty}^{\infty} \bar{x}(t-iT)$$

periodikus jel, tehát Fourier-sora, azaz vonalas spektruma van. Kimondhatjuk, hogy a vizsgált jelben akkor és csak akkor van periodikus összetevő, ha $\bar{x}(t) \neq 0$.

A periodikus összetevő jelenlétével kapcsolatban a következőket lehet mondani.

- Káros, mivel nem hordoz információt, így feleslegesen szállít energiát.
- Hasznos, mivel módot ad arra, hogy a periodikus komponens segítségével helyreállítsuk az órajeleket vagy a vivőfrekvenciás összetevőt (szinkronizáció).

Az

$$s_1(t) = \sum_{i=-\infty}^{\infty} x'_{k_i}(t-iT)$$

folytonos spektrumú, és ez a jel szállítja a hasznos információt.

Az $s_2(t)$ periodikus jel teljesítménysűrűség-függvénye

Az $s_2(t)$ Fourier-sora az alábbi formában írható fel:

$$s_2(t) = \sum_{i=-\infty}^{\infty} \bar{x}(t-iT) = \sum_{l=-\infty}^{\infty} c_l e^{j\omega_0 l t},$$

ahol

$$c_l = \frac{1}{T} \int_0^T s_2(t) e^{-j\omega_0 l t} dt, \quad \omega_0 = \frac{2\pi}{T}.$$

A c_l számítása egyszerű, mivel

$$\begin{aligned}
 c_l &= \frac{1}{T} \int_0^T \sum_{i=-\infty}^{\infty} \bar{x}(t-iT) e^{-j\omega_0 l t} dt = \\
 &= \sum_{i=-\infty}^{\infty} \frac{1}{T} \int_0^T \bar{x}(t-iT) e^{-j\omega_0 l t} dt = \\
 &= \sum_{i=-\infty}^{\infty} \frac{1}{T} \int_{-iT}^{T-iT} \bar{x}(\rho) e^{-j\omega_0 l (\rho+iT)} d\rho = \\
 &= \frac{1}{T} \int_{-\infty}^{\infty} \bar{x}(\rho) e^{-j\omega_0 l \rho} d\rho = \\
 &= \frac{1}{T} \bar{X}(l\omega_0) = \\
 &= \frac{1}{T} \sum_{k=1}^M \pi_k X_k(l\omega_0),
 \end{aligned}$$

mivel

$$e^{-j\omega_0 l iT} = e^{-j \frac{2\pi}{T} l iT} = e^{-j l i 2\pi} = 1$$

és

$$\bar{X}(\omega) = \mathcal{F}\{\bar{x}(t)\}, \quad X_k(\omega) = \mathcal{F}\{x_k(t)\}.$$

$$\begin{aligned}
 R_{s_2}(\tau) &= \frac{1}{T} \int_0^T s_2(t) s_2(t+\tau) dt = \\
 &= \frac{1}{T} \int_0^T \left(\sum_{l=-\infty}^{\infty} c_l e^{j\omega_0 l t} \right) \left(\sum_{h=-\infty}^{\infty} c_h e^{j\omega_0 h (t+\tau)} \right) dt = \\
 &= \frac{1}{T} \int_0^T \sum_{l=-\infty}^{\infty} \sum_{h=-\infty}^{\infty} c_l c_h e^{j\omega_0 (l+h)t} e^{j\omega_0 h \tau} dt \stackrel{l=-h}{=} \\
 &= \frac{1}{T} \int_0^T \sum_{l=-\infty}^{\infty} |c_l|^2 e^{j\omega_0 l \tau} dt = \\
 &= \sum_{l=-\infty}^{\infty} |c_l|^2 e^{j\omega_0 l \tau},
 \end{aligned}$$

mivel

$$\frac{1}{T} \int_0^T e^{j\omega_0 (l+h)t} dt = \begin{cases} 1, & l = -h \\ 0, & l \neq -h \end{cases}$$

Ezek alapján az $s_2(t)$ spektruma a következő alakban adható meg:

$$s_{s_2}(\omega) = \sum_{l=-\infty}^{\infty} |c_l|^2 \delta(\omega - l\omega_0) = \frac{1}{T^2} \sum_{l=-\infty}^{\infty} \left| \sum_{k=1}^M \pi_k X_k(l\omega_0) \right|^2 \delta(\omega - l\omega_0),$$

mivel

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega_0 t \tau} e^{-j\omega \tau} d\tau = \delta(\omega - \omega_0 t).$$

Az $s_1(t)$ teljesítménysűrűség-függvénye statisztikailag független elemi jelsorozat esetén

$$s_1(t) = \sum_{i=-\infty}^{\infty} x'_{k_i}(t - iT)$$

Tételezzük fel, hogy k_i és k_j független valószínűségi változók, ha $i \neq j$, emellett tudjuk, hogy $\mathbf{E}\{x'_k(\tau)\} = 0$.

$$\begin{aligned} R_{s_1}(\tau) &= \frac{1}{T} \int_0^T \mathbf{E}\{s_1(t)s_1(t+\tau)\} dt = \\ &= \frac{1}{T} \int_0^T \mathbf{E}\left\{ \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} x'_{k_i}(t-iT)x'_{k_j}(t+\tau-jT) \right\} dt = \\ &= \frac{1}{T} \int_0^T \sum_{i=-\infty}^{\infty} \mathbf{E}\{x'_{k_i}(t-iT)x'_{k_i}(t+\tau-iT)\} dt = \\ &= \sum_{i=-\infty}^{\infty} \frac{1}{T} \int_{-iT}^{T-iT} \mathbf{E}\{x'_{k_0}(t)x'_{k_0}(t+\tau)\} dt = \\ &= \frac{1}{T} \int_{-\infty}^{\infty} \mathbf{E}\{x'_{k_0}(t)x'_{k_0}(t+\tau)\} dt = \\ &= \frac{1}{T} \int_{-\infty}^{\infty} \sum_{k=1}^M \pi_k x'_k(t)x'_k(t+\tau) dt = \\ &= \sum_{k=1}^M \pi_k \frac{1}{T} \int_{-\infty}^{\infty} x'_k(t)x'_k(t+\tau) dt \end{aligned}$$

A PAM jel elemzése alapján a spektrális sűrűség egyszerűen adódik, mivel láttuk, hogy

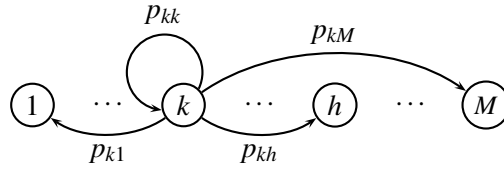
$$R(\tau) = \frac{1}{T} \int_{-\infty}^{\infty} g_T(t)g_T(t+\tau) dt$$

esetén

$$s(\omega) = \frac{1}{2\pi T} |G_T(\omega)|^2,$$

így esetünkben

$$\begin{aligned} s_{s_1}(\omega) &= \frac{1}{2\pi T} \sum_{k=1}^M \pi_k |X'_k(\omega)|^2 = \\ &= \frac{1}{2\pi T} \sum_{k=1}^M \pi_k |X_k(\omega) - \bar{X}(\omega)|^2 = \end{aligned}$$



8.12. ábra. Az elemi jelek sorozata által alkotott Markov-lánc

$$\begin{aligned}
 &= \frac{1}{2\pi T} \sum_{k=1}^M \pi_k (X_k(\omega) - \bar{X}(\omega)) (X_k(\omega) - \bar{X}(\omega))^* = \\
 &= \frac{1}{2\pi T} \sum_{k=1}^M \pi_k |X_k(\omega)|^2 - \frac{1}{2\pi T} |\bar{X}(\omega)|^2
 \end{aligned}$$

A teljes $s(t)$ jel spektrumára tehát az

$$s_s(\omega) = \frac{1}{T^2} \sum_{l=-\infty}^{\infty} \left| \sum_{k=1}^M \pi_k X_k(l\omega_0) \right|^2 \delta(\omega - l\omega_0) + \frac{1}{2\pi T} \sum_{k=1}^M \pi_k |X_k(\omega)|^2 - \frac{1}{2\pi T} |\bar{X}(\omega)|^2$$

kifejezést kapjuk. Nyilvánvaló, hogy $\bar{x}(t) = 0$ esetén $\bar{X}(\omega) = 0$, így

$$s_s(\omega) = \frac{1}{2\pi T} \sum_{k=1}^M \pi_k |X_k(\omega)|^2.$$

Az $s_1(t)$ teljesítménysűrűség-függvénye statisztikailag nem független elemi jelsorozat esetén

Az elemi jelek sorozatának, azaz a $\{k_i\}$ sorozat elemeinek a statisztikai függősége a modulált jelek teljesítménysűrűség-függvényét befolyásolhatja. E függőséget azért is létrehozhatjuk, hogy a jel spektrumát ezáltal illesszük a csatornához.

Tételezzük fel, hogy a $\{k_i\}$ sorozat Markov-láncot képez a

$$p_{kh} = p_{kh}^{(1)} = \mathbf{P}(k_{i+1} = h \mid k_i = k)$$

állapotátmenet-valószínűségekkel, és a rendszer k -adik állapota a 8.12. ábrán azt jelenti, hogy a vizsgált időrészben a rendszer éppen a k -adik elemi jelet viszi át.

Jelöljük $\mathbf{\Pi} = \{p_{kh}\}$ -val a rendszer állapotátmenet-mátrixát, és tudjuk, hogy

$$\sum_{h=1}^M p_{kh} = 1.$$

Ha $\boldsymbol{\pi}_i$ -vel jelöljük az állapotvalószínűségek vektorát az i -edik időrészben, akkor

$$\boldsymbol{\pi}_i = \boldsymbol{\pi}_{i-1} \mathbf{\Pi} = \boldsymbol{\pi}_{i-2} \mathbf{\Pi}^2 = \dots = \boldsymbol{\pi}_0 \mathbf{\Pi}^i,$$

$$\boldsymbol{\pi}_i = (\pi_{i1}, \dots, \pi_{iM})$$

és

$$\pi_{ih} = \sum_{k=1}^M \pi_{(i-1)k} p_{kh}.$$

Ily módon

$$\boldsymbol{\pi}_n = \boldsymbol{\pi}_0 \mathbf{\Pi}^n$$

ahol $\mathbf{\Pi}^n$ az n lépéses átmenetvalószínűségek mátrixa, azaz

$$\mathbf{\Pi}^n = \{p_{kh}^{(n)}\}, \quad p_{kh}^{(n)} = \mathbf{P}(k_{i+n} = k \mid k_i = k).$$

Ha a lépések száma minden határon túl nő, akkor eljutunk a

$$\boldsymbol{\pi} = \boldsymbol{\pi}_0 \mathbf{\Pi}^\infty = \boldsymbol{\pi} \mathbf{\Pi}, \quad \boldsymbol{\pi} = (\pi_1, \dots, \pi_M)$$

stationer állapotvalószínűség-vektorhoz, vagyis az elemi jelek a priori valószínűségeinek vektorához.

Az $s_1(t)$ jel spektrumának számítása most valamivel bonyolultabb, mint az előző esetekben.

$$s_1(t) = \sum_{i=-\infty}^{\infty} x'_{k_i}(t - iT),$$

de most tételezzük fel, hogy $\bar{x}(t) = 0$, ezért $x'(t) = x(t)$, így

$$s_1(t) = \sum_{i=-\infty}^{\infty} x_{k_i}(t - iT).$$

$$R_{s_1}(\tau) = \frac{1}{T} \int_0^T \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \mathbf{E} \{x_{k_i}(t - iT)x_{k_j}(t + \tau - jT)\} dt,$$

de most a várható érték képzéskor figyelembe kell venni a különböző $\{k_i\}$ indexek függőségét is:

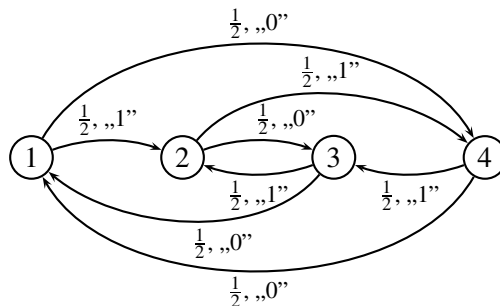
$$\mathbf{E} \{x_{k_i}(t - iT)x_{k_j}(t + \tau - jT)\} = \begin{cases} \sum_{k=1}^M \pi_k x_k(t - iT)x_k(t + \tau - iT), & i = j \\ \sum_{k=1}^M \sum_{h=1}^M \pi_k p_{kh}^{(j-i)} x_k(t - iT)x_h(t + \tau - jT), & j > i \\ \sum_{k=1}^M \sum_{h=1}^M \pi_k p_{kh}^{(i-j)} x_h(t - iT)x_k(t + \tau - jT), & j < i \end{cases}$$

Ebből a korábbiakhoz hasonlóan

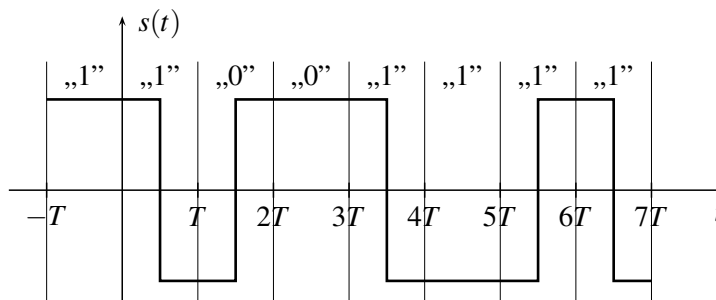
$$\begin{aligned} R_{s_1}(\tau) &= \frac{1}{T} \int_{-\infty}^{\infty} \sum_{k=1}^M \pi_k x_k(t)x_k(t + \tau) dt + \\ &+ \frac{1}{T} \int_{-\infty}^{\infty} \sum_{l=1}^{\infty} \sum_{k=1}^M \sum_{h=1}^M \pi_k p_{kh}^{(l)} (x_k(t)x_h(t + \tau - lT) + x_h(t)x_k(t + \tau + lT)) dt, \end{aligned}$$

amiből

$$\begin{aligned} s_{s_1}(\omega) &= \frac{1}{2\pi T} \sum_{k=1}^M \pi_k |X_k(\omega)|^2 + \\ &+ \frac{1}{2\pi T} \sum_{k=1}^M \sum_{h=1}^M \pi_k \sum_{l=1}^{\infty} p_{kh}^{(l)} \left(e^{-j\omega l T} X_k^*(\omega) X_h(\omega) + e^{j\omega l T} X_k(\omega) X_h^*(\omega) \right) \end{aligned}$$



8.13. ábra. A Miller-kód állapotátmenet-diagramja



8.14. ábra. A Miller-kódolt jel egy realizációja

Példa a nem független elemi jelek esetére (Miller-kód)

Feltételek:
Elemi jelek

$$x_1(t) = -x_4(t) = \begin{cases} \sqrt{\frac{E}{T}}, & t \in [0, T) \\ 0, & t \notin [0, T) \end{cases}$$

$$x_2(t) = -x_3(t) = \begin{cases} \sqrt{\frac{E}{T}}, & t \in [0, \frac{T}{2}) \\ -\sqrt{\frac{E}{T}}, & t \in [\frac{T}{2}, T) \\ 0, & t \notin [0, T) \end{cases}$$

Bináris átvitelt használunk, azaz az állapotátmeneti mátrix egy-egy sorában csak két nullától különböző érték van. A bináris forrásszimbólumok azonos eséllyel generálódnak, és

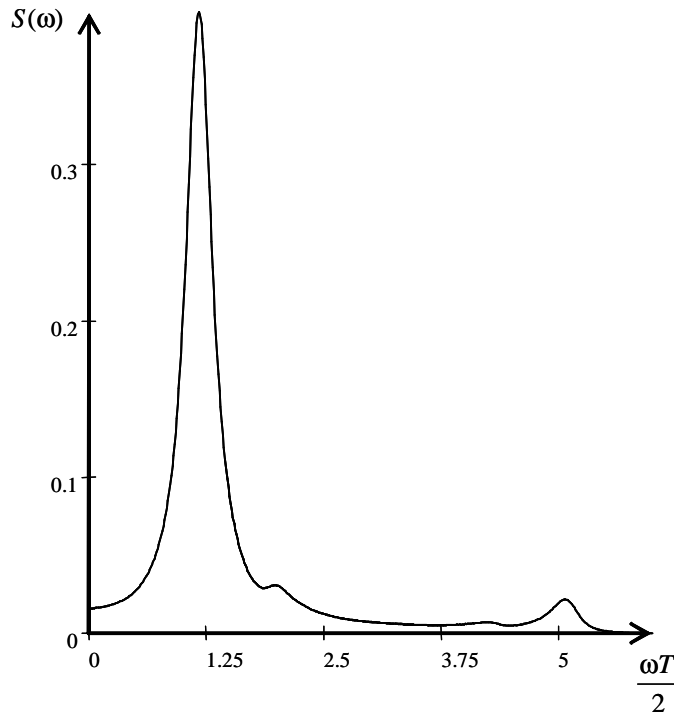
$$\mathbf{\Pi} = \begin{bmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{bmatrix}$$

Az állapotátmenet-diagram a 8.13. ábrán látható.

A modulált jel egy realizációját a 8.14. ábrán adtuk meg. Amiből jól érzékelhető, hogy a modulált jel sávzélessége kisebb, mint a Manchester-kódolt jelé, és a spektrumban kicsi a DC komponens szintje. Ezen kívül $\boldsymbol{\pi} = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$, mivel

$$\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{\Pi} = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}) \begin{bmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{bmatrix} = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$$

A jel spektrálisan jól illeszkedik a mágneses jelrögzítés csatornájához, mivel a domináns spektrális komponensek a $\frac{\pi}{2T}$ közelében találhatók, azaz a jel sávszélessége viszonylag kicsi.



8.15. ábra. A Miller-kódolt jel teljesítménysűrűség függvénye

Az analitikus eredmények kiértékelése után

$$s(\omega) = \frac{E}{2\pi} \frac{1}{2\Theta^2(17 + 8\cos(8\Theta))} (23 - 2\cos(\Theta) - 22\cos(2\Theta) - 12\cos(3\Theta) + 5\cos(4\Theta) + 12\cos(5\Theta) + 2\cos(6\Theta) - 8\cos(7\Theta) + 2\cos(8\Theta)),$$

ahol $\Theta = \frac{\omega T}{2}$.

A Miller-kódolt jel teljesítménysűrűség függvénye a 8.15. ábrán látható.

8.9. A folytonos fázisú FM modulált jelek spektruma

(CPM, Continuous Phase Modulation)

Feltételek:

- Legyen $\{\xi_{k_i}\}$ független sorozat.
- $\xi_{k_i} \in \{-(M-1), \dots, -3, -1, 1, 3, \dots, (M-1)\}$
- $\pi_k = \mathbf{P}(\xi_{k_i} = \xi_k)$
- $g(t)$ az elemi frekvencia-időfüggvény, melynek tartója a $[0, LT)$ tartomány.
- $q(t) = \int_{-\infty}^t g(\vartheta) d\vartheta$ az elemi fázisfüggvény
- $q(LT) = \frac{1}{2}$ (normált érték, elvileg lehet más is, pl. nulla)

A modulált jel alakja

$$s(t) = \sqrt{2P} \cos(\omega_0 t + \Phi(t, \xi) + \varphi),$$

ahol φ a vivő véletlen fázisa, egy független, egyenletes eloszlású valószínűségi változó a $[0, 2\pi)$ tartományban, $\Phi(t, \xi)$ az információt hordozó fázis-időfüggvény és P a jel teljesítménye.

A jel korrelációs függvénye az

$$r(\tau) = 2P \mathbf{E}_\varphi \{ \mathbf{E}_t \{ \mathbf{E}_\xi \{ s(t)s(t+\tau) \} \} \}$$

kifejezéssel adható meg, ahol

$$\mathbf{E}_\varphi \{ \cdot \} = \frac{1}{2\pi} \int_0^{2\pi} (\cdot) d\varphi$$

és

$$\mathbf{E}_t \{ \cdot \} = \frac{1}{T} \int_0^T (\cdot) dt.$$

$$\begin{aligned} r(\tau) &= 2P \mathbf{E}_{\varphi,t,\xi} \{ \cos(\omega_0 t + \Phi(t, \xi) + \varphi) \cos(\omega_0(t+\tau) + \Phi(t+\tau, \xi) + \varphi) \} = \\ &= P \mathbf{E}_{\varphi,t,\xi} \{ \cos(\omega_0 \tau + \Phi(t+\tau, \xi) - \Phi(t, \xi)) + \cos(\omega_0(2t+\tau) + \Phi(t+\tau, \xi) + \Phi(t, \xi) + 2\varphi) \} = \\ &= P \mathbf{E}_{t,\xi} \{ \cos(\omega_0 \tau + \Phi(t+\tau, \xi) - \Phi(t, \xi)) \} = \\ &= P \mathbf{E}_{t,\xi} \{ \operatorname{Re} \{ \exp(j(\Phi(t+\tau, \xi) - \Phi(t, \xi))) \exp(j\omega_0 \tau) \} \} = \\ &= P \operatorname{Re} \{ \exp(j\omega_0 \tau) \mathbf{E}_{t,\xi} \{ \exp(j(\Phi(t+\tau, \xi) - \Phi(t, \xi))) \} \} \end{aligned}$$

Ebben a kifejezésben P a jel teljesítménye, $\exp(j\omega_0 \tau)$ a vivőfrekvenciás összetevő, $R(\tau)$ pedig az alapsávi komplex korrelációs függvény.

A spektrum számításához elegendő az $R(\tau)$ vizsgálata.

$$R(\tau) = \mathbf{E}_{t,\xi} \{ \exp(j(\Phi(t+\tau, \xi) - \Phi(t, \xi))) \},$$

ahol

$$\Phi(t, \xi) = 2\pi h \sum_{i=-\infty}^{\infty} \xi_{k_i} q(t - iT),$$

ezért

$$\begin{aligned} R(\tau) &= \mathbf{E}_{t,\xi} \left\{ \exp \left(j2\pi h \sum_{i=-\infty}^{\infty} \xi_{k_i} (q(t+\tau - iT) - q(t - iT)) \right) \right\} = \\ &= \mathbf{E}_{t,\xi} \left\{ \prod_{i=-\infty}^{\infty} \exp(j2\pi h \xi_{k_i} (q(t+\tau - iT) - q(t - iT))) \right\}. \end{aligned}$$

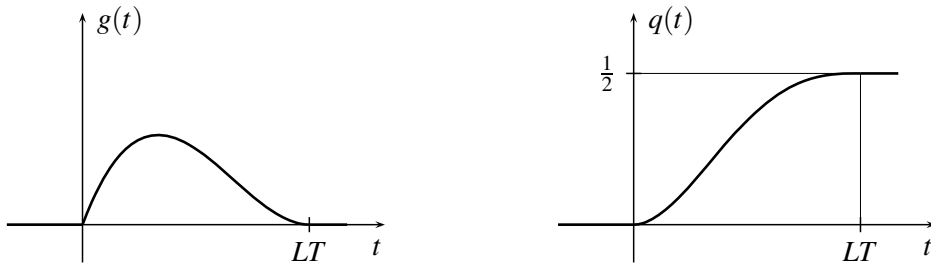
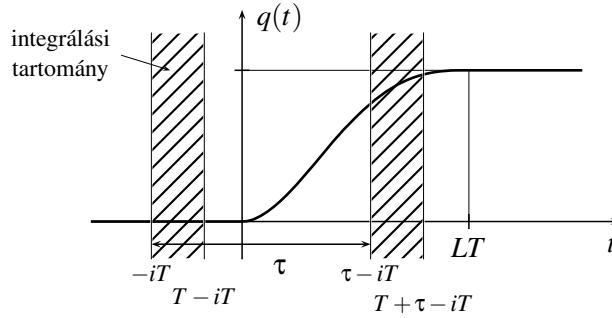
Ha a ξ_{k_i} független sorozat, akkor a várható értéket tagonként lehet előállítani, azaz

$$R(\tau) = \frac{1}{T} \int_0^T \prod_{i=-\infty}^{\infty} \mathbf{E}_{\xi_{k_i}} \{ \exp(j2\pi h \xi_{k_i} (q(t+\tau - iT) - q(t - iT))) \} dt.$$

A 8.16. ábrán a $g(t)$ és $q(t)$ függvényeket illusztráljuk, ha $g(t)$ tartója LT .

A továbbiakban az $R(\tau)$ kiszámításával foglalkozunk.

$$\begin{aligned} R(\tau) &= \frac{1}{T} \int_0^T \prod_{i=-\infty}^{\infty} \mathbf{E}_{\xi_{k_i}} \{ \exp(j2\pi h \xi_{k_i} (q(t+\tau - iT) - q(t - iT))) \} dt = \\ &= \frac{1}{T} \int_0^T \prod_{i=-\infty}^{\infty} \left(\sum_{\substack{k=-(M-1) \\ \text{páratlan}}}^{M-1} \pi_k \exp(j2\pi h k (q(t+\tau - iT) - q(t - iT))) \right) dt. \end{aligned}$$

8.16. ábra. A $g(t)$ és $q(t)$ függvények8.17. ábra. A $q(t)$ függvény vizsgálata

Az elvégzendő műveletek illusztrálására szolgál a 8.17. ábra. Az illusztráció alapján világos, hogy rögzített $\tau > 0$ mellett csak véges szorzatot kell figyelembe venni, mivel, ha $t - iT > LT$ vagy $t + \tau - iT + T < 0$, akkor

$$q(t + \tau - iT) - q(t - iT) = 0,$$

így a szorzótényező

$$\exp(j2\pi hk(q(t + \tau - iT) - q(t - iT))) = 1.$$

Számítsuk ki $R(\tau)$ értékét $\tau > LT$ értékeknél oly módon, hogy $\tau = mL + \tau'$, $\tau' \in [0, T)$ és $m > L$.

- Legyen először

$$\begin{aligned} 0 &> t + mT + \tau' - iT \\ iT &> t + mT + \tau' \\ iT &> (m+2)T, \end{aligned}$$

ennek inverze

$$\begin{aligned} iT &\leq (m+1)T \\ i &\leq m+1 \end{aligned}$$

- Vizsgáljuk meg másodszor

$$\begin{aligned} LT &< t - iT \\ iT &< t - LT, \end{aligned}$$

ennek inverze

$$\begin{aligned} iT &\geq -LT + T \\ i &\geq 1 - L \end{aligned}$$

Ezért az $R(\tau)$ az alábbi módon számítható:

$$R(\tau) = \frac{1}{T} \int_0^T \prod_{i=1-L}^{m+1} \left(\sum_{\substack{k=-(M-1) \\ \text{páratlan}}}^{M-1} \pi_k \exp(j2\pi hk(q(t+\tau-iT) - q(t-iT))) \right) dt.$$

A $q(t+\tau-iT) - q(t-iT)$ különbség értékei három csoportba sorolhatók

- $q(t+\tau-iT) - q(t-iT) = q(t+\tau-iT)$, ha $q(t-iT) = 0$ és

$$\begin{aligned} 0 \leq t + \tau - iT &\leq LT && (m-i)T \geq -(t+\tau') \\ t + mT + \tau' - iT &\leq LT && (m-i)T \geq -T \\ (m-i)T &\leq LT - (t+\tau') && \\ (m-i)T &\leq LT - T && \\ m-i &\leq L-1 && m-i \geq -1 \end{aligned}$$

- $q(t+\tau-iT) - q(t-iT) = q(LT) - q(t-iT) = \frac{1}{2} - q(t-iT)$, ha $q(t+\tau-iT) = q(LT)$ és

$$\begin{aligned} 0 \leq t - iT &\leq LT && iT \leq t \\ iT &\geq -LT + t && iT \leq 0 \\ iT &\geq -LT + T && iT \leq 0 \\ i &\geq 1-L && i \leq 0 \end{aligned}$$

- $q(t+\tau-iT) - q(t-iT) = q(LT)$, ha $q(t-iT) = 0$, $t+\tau-iT > LT$ és $t-iT < 0$. Az összes nem semleges szorzatok száma

$$m+1 - (1-L) + 1 = m+L+1.$$

A $q(t+\tau-iT)$ értékek száma

$$L-1 - (-1) + 1 = L+1.$$

A $q(LT) - q(t-iT)$ értékek száma

$$0 - (1-L) + 1 = L.$$

Ezért a $q(LT)$ értékek száma

$$m+L+1 - (L+1) - L = m-L$$

A komplex alapsávi korrelációs függvényt az alábbi módon számíthatjuk:

$$\begin{aligned} R(\tau) &= \frac{1}{T} \int_0^T \prod_{i=1-L}^0 \left(\sum_{\substack{k=-(M-1) \\ \text{páratlan}}}^{M-1} \pi_k \exp(j2\pi hk(q(LT) - q(t-iT))) \right) \cdot \\ &\quad \prod_{i=1-L}^1 \left(\sum_{\substack{k=-(M-1) \\ \text{páratlan}}}^{M-1} \pi_k \exp(j2\pi hk(q(t+\tau'-iT))) \right) \cdot \\ &\quad \left(\sum_{\substack{k=-(M-1) \\ \text{páratlan}}}^{M-1} \pi_k \exp(j2\pi hkq(LT)) \right)^{m-L} dt. \end{aligned}$$

Ennek alapján belátható, hogy

$$R(\tau) = C_{\xi}^{m-L} \Psi(\tau') = R(\tau' + mT),$$

ahol

$$C_{\xi} = \sum_{\substack{k=-(M-1) \\ \text{páratlan}}}^{M-1} \pi_k \exp(j2\pi h k q(LT))$$

konstans.

A jel teljesítménysűrűsége az $R(\tau)$ Fourier-transzformáltjából számolható. Vegyük észre, hogy $R(\tau)$ két függvény szorzataként állítható elő, ha $\tau = mT + \tau' > LT$.

A korábbiakból tudjuk, hogy modulált jel korrelációs függvénye

$$r(\tau) = P \operatorname{Re} \{ R(\tau) \exp(j\omega_0 \tau) \}.$$

Legyen a továbbiakban $P = 1$, és írjuk fel a jel teljesítményspektrumát

$$\begin{aligned} s(\omega) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \operatorname{Re} \{ R(\tau) \exp(j\omega_0 \tau) \} e^{-j\omega \tau} d\tau = \\ &= \frac{1}{2\pi} \int_0^{\infty} \operatorname{Re} \{ R(\tau) \exp(j\omega_0 \tau) \} e^{-j\omega \tau} d\tau + \\ &\quad + \frac{1}{2\pi} \int_{-\infty}^0 \operatorname{Re} \{ R(\tau) \exp(j\omega_0 \tau) \} e^{-j\omega \tau} d\tau = \\ &= \frac{2}{2\pi} \operatorname{Re} \left\{ \int_0^{\infty} \operatorname{Re} \{ R(\tau) \exp(j\omega_0 \tau) \} e^{-j\omega \tau} d\tau \right\} = \\ &= \frac{2}{2\pi} \operatorname{Re} \left\{ \int_0^{\infty} \frac{R(\tau) \exp(j\omega_0 \tau) + R^*(\tau) \exp(-j\omega_0 \tau)}{2} e^{-j\omega \tau} d\tau \right\} = \\ &= \frac{1}{2\pi} \operatorname{Re} \left\{ \int_0^{\infty} R(\tau) e^{-j(\omega - \omega_0)\tau} d\tau + \int_0^{\infty} R^*(\tau) e^{-j(\omega + \omega_0)\tau} d\tau \right\}. \end{aligned}$$

Ebből nyilvánvaló, hogy elegendő az

$$S(\omega) = \frac{1}{2\pi} \operatorname{Re} \left\{ \int_0^{\infty} R(\tau) e^{-j\omega \tau} d\tau \right\}$$

meghatározása.

A kijelölt integrál a következőképpen számolható:

$$\int_0^{\infty} R(\tau) e^{-j\omega \tau} d\tau = \int_0^{LT} R(\tau) e^{-j\omega \tau} d\tau + \int_{LT}^{\infty} R(\tau) e^{-j\omega \tau} d\tau,$$

ahol a második tag

$$\int_{LT}^{\infty} R(\tau) e^{-j\omega \tau} d\tau = \sum_{m=L}^{\infty} \int_0^T R(mT + \tau') e^{-j\omega(mT + \tau')} d\tau' =$$

$$\begin{aligned}
&= \sum_{m=L}^{\infty} \int_0^T C_{\xi}^{m-L} \Psi(\tau') e^{-j\omega(mT+\tau')} d\tau' = \\
&= \sum_{m=L}^{\infty} C_{\xi}^{m-L} e^{-j\omega mT} \int_0^T \Psi(\tau') e^{-j\omega\tau'} d\tau' = \\
&= \sum_{m=0}^{\infty} C_{\xi}^m e^{-j\omega(m+L)T} \int_0^T \Psi(\tau') e^{-j\omega\tau'} d\tau' = \\
&= \left(\int_0^T \Psi(\tau') e^{-j\omega\tau'} d\tau' \right) e^{-j\omega LT} \sum_{m=0}^{\infty} C_{\xi}^m e^{-j\omega mT},
\end{aligned}$$

és az utolsó tényező egy mértani sor összege, ezért, ha $|C_{\xi}| < 1$, akkor

$$\int_{LT}^{\infty} R(\tau) e^{-j\omega\tau} d\tau = \left(\int_0^T \Psi(\tau') e^{-j\omega\tau'} d\tau' \right) \frac{e^{-j\omega LT}}{1 - C_{\xi} e^{-j\omega T}}.$$

Ezek felhasználásával

$$\begin{aligned}
S(\omega) &= \frac{1}{2\pi} \operatorname{Re} \left\{ \int_0^{\infty} R(\tau) e^{-j\omega\tau} d\tau \right\} = \\
&= \frac{1}{2\pi} \operatorname{Re} \left\{ \int_0^{LT} R(\tau) e^{-j\omega\tau} d\tau + \frac{e^{-j\omega LT}}{1 - C_{\xi} e^{-j\omega T}} \int_0^T \Psi(\tau') e^{-j\omega\tau'} d\tau' \right\}.
\end{aligned}$$

Ha viszont $|C_{\xi}| = 1$, akkor nagy $\tau = mT + \tau' > LT$ esetén az

$$R(\tau) = C_{\xi}^{m-L} \Psi(\tau')$$

nT -re periodikus, ha $C_{\xi}^n = 1$

$$R((m+n)T + \tau') = C_{\xi}^{m+n-L} \Psi(\tau') = C_{\xi}^n R(mT + \tau').$$

Ha viszont a korrelációs függvénynek van τ -ban periodikus összetevője, akkor a teljesítménysűrűség-függvényben vannak vonalas komponensek, azaz a jel tartalmaz periodikus — információt nem hordozó — összetevőt. Ennek alapján a modulált CPM jelnek akkor van periodikus összetevője, ha

$$|C_{\xi}| = \left| \sum_{\substack{k=-(M-1) \\ \text{páratlan}}}^{M-1} \pi_k \exp(j2\pi h k q(LT)) \right| = 1.$$

Vizsgáljuk meg ennek a feltételeit!

- Ha $q(LT) = 0$, akkor biztosan igaz, hogy

$$|C_{\xi}| = \sum_{k=-(M-1)}^{M-1} \pi_k = 1,$$

amiből egyértelmű, hogy van periodikus összetevő. Éppen ez az oka annak, hogy $q(LT) = \frac{1}{2} \neq 0$ értékeket használunk csak.

- Ha $q(LT) \neq 0$, akkor $|C_\xi|$ csak oly módon lehet 1, ha

$$\exp(j2\pi hkq(LT)) = \exp(jb)$$

minden k -ra azonos értékű, mivel ilyenkor

$$C_\xi = \sum_{\substack{k=-(M-1) \\ \text{páratlan}}}^{M-1} \pi_k \exp(j2\pi hkq(LT)) = \exp(jb) \sum_{\substack{k=-(M-1) \\ \text{páratlan}}}^{M-1} \pi_k = \exp(jb).$$

A periodikus komponens jelenlétének tehát $q(LT) = \frac{1}{2}$ esetén az a feltétele, hogy teljesüljenek a következők:

$$\begin{aligned} \pi hk &= b \pmod{2\pi}, \\ \pi h(k+2) &= b \pmod{2\pi}, \\ 2\pi h &= 0 \pmod{2\pi}, \end{aligned}$$

tehát h legyen egész szám.

1. Ha $C_\xi = 1$, $h = 0, \pm 2, \pm 4, \dots$, akkor $C_\xi^1 = 1$, így

$$R(\tau) = R(mT + \tau') = C_\xi^{m-L} \Psi(\tau')$$

T -re periodikus, vagyis a periodikus összetevők frekvenciája a $\frac{2\pi}{T}$ egész számú többszöröse.

2. Ha $C_\xi = -1$, $h = \pm 1, \pm 3, \dots$, akkor $C_\xi^2 = 1$, így

$$R(\tau) = C_\xi^{m-L} \Psi(\tau')$$

$2T$ -re periodikus, vagyis a periodikus összetevők frekvenciája a $\frac{2\pi}{2T}$ egész számú többszöröse.