

Moduláció

Bevezetés

A moduláció lényege, célja. A *moduláció* röviden változtatást jelent. Például a zenében a kulcs megváltoztatását, míg a hírközlésben a vivőhullám valamely jellemzőjének változtatását nevezik modulációnak. Persze ez rögtön felveti azt a kérdést, hogy miért kell egyáltalán vivőhullám? Erre nagyon egyszerű a magyarázat: a továbbítandó, az információt hordozó jelek elég általánosan az alapsávban vannak, viszont a rendelkezésre álló átviteli csatornák igen gyakran ettől eltérő frekvenciasávot biztosítanak. Így a moduláció egyik legalapvetőbb céljaként ennek az eltérésnek a feloldását jelölhetjük meg. Jól ismert, hogy valóban létezik olyan modulációs eljárás, az SSB, amely tényleg csupán ezt a *frekvencia-áthelyezést* valósítja meg, de más modulációs módszerek további lehetőségeket is biztosíthatnak. Ezek a lehetőségek a modulált jel által elfoglalt frekvenciasáv értékében, a modulált jelhez adódó zaj tolerálásában jelentkeznek.

Modulált jelek matematikai modelljei. A modulált jelek leírására a legkézenfekvőbb az időtartomány, de sok esetben itt rejtve maradnak fontos tulajdonságok. Éppen ezért gyakran használjuk a frekvencia-tartományt is, amikor meghatározzuk a modulált jel spektrumát. Digitális moduláció esetén nagyon fontos lesz a *jeltér*-beli leírás, amivel itt fogunk megismerkedni.

A moduláció fajtái. Két alapvető fajtát használunk: az egyik az analóg moduláció, a másik a digitális moduláció.

Az analóg moduláció a legáltalánosabban az $\eta_t = f(\xi_t)$ leképezést jelenti, ahol a ξ_t moduláló jelet az $f(\cdot)$ funkcionál leképezi az η_t függvénybe. A leggyakrabban használt analóg modulációs eljárások esetén egy szinuszos vivőhullám amplitúdóját, illetve fázisát vagy frekvenciáját változtatja, azaz modulálja a ξ_t jel.

Digitális modulációnál a moduláló szimbólumokat leképezzük a csatornán továbbítható „szimbólumokká”, más néven *elemi jelek*-ké. Ebben az esetben ugyanis nincs eleve adott jel, amit *alakhűen* kellene továbbítani. A modulált szimbólumok célszerű leírása a jeltérben történhet.

A demoduláció. Analóg esetben a moduláló jelet akarjuk alakhűen visszaállítani. Digitális esetben a csatorna kimenetén lévő jelből el akarjuk dönteni a legvalószínűbb küldött szimbólumot. Akár azt is mondhatjuk, hogy nincs demoduláció, csak döntés van.

A demoduláció feladata persze nagymértékben különbözik zajmentes esetben, illetve zajos esetben. (Valóságos feladat általában az utóbbi esetben fordul elő.) A zajmentes esetben az analóg demoduláció a modulációs leképezés inverz leképezésének a meghatározását jelenti csupán. Persze zajos esetben is valami hasonlóra van szükség, de inkább úgy fogalmazhatunk, hogy meg kell határozni a zajos modulált jelből a moduláló jelhez valamilyen értelemben legközelebb álló jelet.

Digitális esetben a demoduláció a csatorna szimbólumok leképezése a moduláló szimbólumokba. Természetesen ez a leképezés zajos esetben igazán érdekes feladat.

A demodulációval kapcsolatban lényegében az a két kérdés vetődik fel, hogy: Mi a legjobb eljárás? Mi a legjobb eredmény?

Analóg modulációs eljárások

Ide jön a régi 6-os

Digitális modulációs eljárások

Mint mondtuk, a digitális moduláció a forrás-szimbólumoknak a csatorna-szimbólumokba történő leképezését jelenti. A csatorna-szimbólumok, azaz a digitálisan modulált jel elemeinek a leírására a jeltér a legalkalmasabb.

A jeltér

Céljainknak megfelelően azt vizsgáljuk, hogy miként lehet a véges energiájú, korlátozott időtartamú jeleket ($x_i(t), 0 \leq t < \tau_0, i = 1, \dots, M$) egyszerűen, és a moduláció és demoduláció szempontjából hatékonyan leírni. Az derül ki, hogy egy ortonormált vektortérben, a *jeltérben* történő leírás rendkívül előnyös.

A vektorteret az alábbi jellemzőkkel bíró bázisfüggvényekkel feszítjük ki:

$$\varphi_j(t), 0 \leq t < \tau_0, j = 1, \dots, N,$$

úgy, hogy skaláris szorzatuk:

$$\langle \varphi_j(t), \varphi_k(t) \rangle = \int_0^{\tau_0} \varphi_j(t) \cdot \varphi_k(t) dt = \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases}$$

Azaz ortonormált a tér, mert a bázisfüggvényei ortogonálisak, és energiájuk egységnyi.

Továbbá úgy választjuk a bázisfüggvényeket, hogy a tér *teljes* legyen:

$$x_i(t) = \sum_{j=1}^N x_{ij} \cdot \varphi_j(t), \quad i = 1, \dots, M; \quad \text{és} \quad x_{ij} = \int_0^{\tau_0} x_i(t) \cdot \varphi_j(t) dt,$$

azaz, a leírni kívánt jeleket, az ú.n. csatorna-szimbólumokat maradéktalanul meghatározzák a bázisfüggvények.

Csak megemlítjük, hogy a fenti követelményeket kielégítő bázisfüggvények meghatározására algoritmikus eljárás ismert, az úgy nevezett Gram-Schmidt ortogonalizáció.

Az előbbi két egyenletből az elsőt szokták szintézis-egyenletnek hívni:

$$x_i(t) = \sum_{j=1}^N x_{ij} \cdot \varphi_j(t), \quad i = 1, \dots, M,$$

a másodikat pedig az analízis egyenletnek is:

$$x_{ij} = \int_0^{\tau_0} x_i(t) \cdot \varphi_j(t) dt; \quad i = 1, \dots, M; \quad \text{és} \quad j = 1, \dots, N.$$

Az egyenletekben szereplő jelölések, amelyeket sokat fogunk használni: M jelöli a digitális szimbólumkészlet méretét, azt is mondjuk, hogy M -áris a forrás; ennek a szimbólumkészletnek megfelelő csatorna-szimbólumoknak, azaz az M -áris modulált szimbólumkészletnek a jeltérbeli leírásához szükséges bázisfüggvények számát jelöli N . látni fogjuk, hogy a gyakran használt modulációs módszerek esetén N kisebb, vagy akár sokkal kisebb, mint M .

A továbbiakban fontos gyakorlati példákon mutatjuk meg a jeltér használatát.

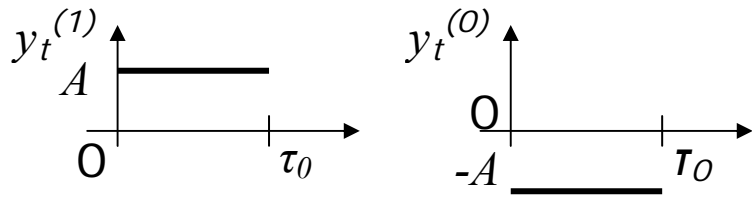
Antipodális jelek

A digitális moduláció legegyszerűbb esetét jelentik az úgy nevezett *antipodális* jelek, amelyet *ellentétes*-nek fordíthatunk, bár ezt az elnevezést nem használjuk. Ezen az egyszerű példán belül is a legegyszerűbb:

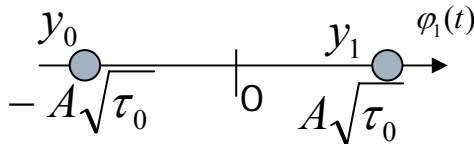
Az antipodális PAM

A bináris szimbólumkészlet értelemszerűen két jeltől áll, amelyek egymás ellentettjei, tükörképei. A τ_0 hosszúságú időrésben a jel változó, vagy konstans értékű is lehet.

Be kell vallani, hogy ebben az esetben a jeltérbeli leírás még bonyolításnak tűnik, hiszen ha a két elemi jel, azaz $y_t^{(1)}$ és $y_t^{(0)}$ egyenlő A és $-A$ az időrés mentén, amint a mellékelt ábra mutatja, akkor ezeknek a leírása az ortonormált jeltér bázisfüggvényei segítségével nem az egyszerűsítés irányába mutat, bár igaz, hogy a jeltér egydimenziós lehet, a következő bázisfüggvénnyel: $\varphi_1(t) = \frac{1}{\sqrt{\tau_0}}, 0 \leq t < \tau_0$, és a jel- „vektorok” (hiszen 1 dimenziósak):



$$y_1 = \frac{y_t^{(1)}}{\varphi_1(t)} = \frac{A}{1/\sqrt{\tau_0}} = A \cdot \sqrt{\tau_0}, \quad \text{és} \quad y_0 = -A \cdot \sqrt{\tau_0}.$$



A modulált jelek jeltérbeli ábrázolása így a mellékelt ábra szerinti. A bázisfüggvény meghatározása csupán annyit jelentett, hogy egységnyi energiájúra normáltuk az 1-es jelű elemi jelet. (Egyébként az említett Gram-Schmidt ortogonalizácónak ez az első lépése.)

$$\int_0^{\tau_0} (k \cdot y_t^{(1)})^2 dt = 1 \Rightarrow \varphi_1(t) = k \cdot y_t^{(1)}$$

A jelvektorok meghatározása már nem szorul magyarázatra.

A most bemutatott példában szereplő NRZ (Non Return to Zero) jel egyszerűsége miatt példaként nagyon előnyös, de gyakorlati alkalmazását tekintve komoly hátrányai vannak. A következő, szintén egyszerű példánk viszont gyakorlatilag is nagy jelentőségű.

A BPSK

A két ellentett jel most:

$$y_t^{(1)} = A \cdot \sin\left(\frac{2\pi}{\tau_0} \cdot t\right), \quad \text{és} \quad y_t^{(0)} = -A \cdot \sin\left(\frac{2\pi}{\tau_0} \cdot t\right); \quad t \in [0, \tau_0).$$

A jeltér ismét egydimenziós lehet, az alábbi bázisfüggvénnyel:

$$\varphi_1(t) = \sqrt{\frac{2}{\tau_0}} \sin\left(\frac{2\pi}{\tau_0} \cdot t\right), \quad 0 \leq t < \tau_0.$$

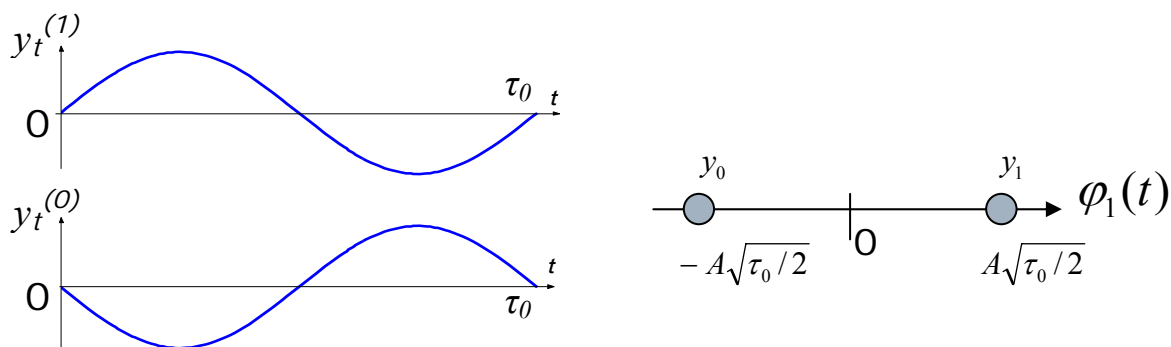
Így a jelvektorok a következők:

$$y_1 = \frac{y_t^{(1)}}{\varphi_1(t)} = A \cdot \sqrt{\frac{\tau_0}{2}}; \quad y_0 = -y_1.$$

A bázisfüggvényt ismét úgy kaptuk, hogy az 1-es jelű elemi jel amplitúdóját a szükséges mértékűre választottuk.

$$\int_0^{\tau_0} (k \cdot y_t^{(1)})^2 dt = \frac{k^2 \cdot A^2 \cdot \tau_0}{2} = 1 \Rightarrow \varphi_1(t) = \frac{\sqrt{2}}{A \sqrt{\tau_0}} \cdot y_t^{(1)}$$

A jelvektorok meghatározása ismét nagyon egyszerű.



Az elemi jeleket és a jelvektorokat illusztrálják az ábrák.

Összehasonlítva a két példa jelvektorait, megállapíthatjuk, hogy az alapvetően eltérő elemi jelek jeltérbeli reprezentációja rendkívül hasonló lehet, és majd látni fogjuk, hogy ebben rejlik a jeltér használatának egyik előnye.

Többállapotú jelek

Az antipodális jelek kétállapotúak, így egy-egy szimbólum maximum 1 bit továbbítását teszi lehetővé. Amennyiben τ_0 idő alatt több információt szeretnénk továbbítani, akkor erre többállapotú jeleket kell használni.

A QPSK

A négy elemi jel legyen a következő:

$$y_t^{(1)} = A \cdot \sin\left(\frac{2\pi}{\tau_0} \cdot t\right) = -y_t^{(0)}, \quad \text{és} \quad y_t^{(2)} = A \cdot \cos\left(\frac{2\pi}{\tau_0} \cdot t\right) = -y_t^{(3)}; \quad t \in [0, \tau_0)$$

A jeltér kétdimenziós lesz, az alábbi két bázisfüggvénnyel, amelyek meghatározása az előzőek szerint történhet, mivel az elemi jelek páronként ortogonálisak:

$$\varphi_1(t) = \sqrt{2/\tau_0} \sin\left(\frac{2\pi}{\tau_0} \cdot t\right); \quad \varphi_2(t) = \sqrt{2/\tau_0} \cos\left(\frac{2\pi}{\tau_0} \cdot t\right), \quad 0 \leq t < \tau_0$$

A jelvektorokat a következő összefüggéssel határozzuk meg:

$$y_{ij} = \int_0^{\tau_0} y_t^{(i)} \cdot \varphi_j(t) dt$$

$$y_{11} = \int_0^{\tau_0} y_t^{(1)} \cdot \varphi_1(t) dt = \int_0^{\tau_0} A \cdot \sin\left(\frac{2\pi}{\tau_0} \cdot t\right) \cdot \sqrt{2/\tau_0} \sin\left(\frac{2\pi}{\tau_0} \cdot t\right) dt = \frac{A\sqrt{2}}{\sqrt{\tau_0}} \cdot \frac{\tau_0}{2} = \frac{A\sqrt{\tau_0}}{\sqrt{2}},$$

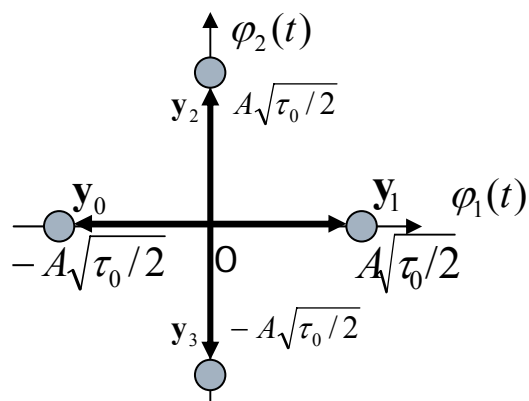
$$y_{12} = \int_0^{\tau_0} y_t^{(1)} \cdot \varphi_2(t) dt = \int_0^{\tau_0} A \cdot \sin\left(\frac{2\pi}{\tau_0} \cdot t\right) \cdot \sqrt{2/\tau_0} \cos\left(\frac{2\pi}{\tau_0} \cdot t\right) dt = 0.$$

Így folytatva a többivel, a következőket kapjuk:

$$\mathbf{y}_1 = \left(\frac{A\sqrt{\tau_0}}{\sqrt{2}}, 0\right); \quad \mathbf{y}_0 = \left(-\frac{A\sqrt{\tau_0}}{\sqrt{2}}, 0\right);$$

$$\mathbf{y}_2 = \left(0, \frac{A\sqrt{\tau_0}}{\sqrt{2}}\right); \quad \mathbf{y}_3 = \left(0, -\frac{A\sqrt{\tau_0}}{\sqrt{2}}\right).$$

A mellékelt ábra illusztrálja a kapott jelvektorokat.



A QAM

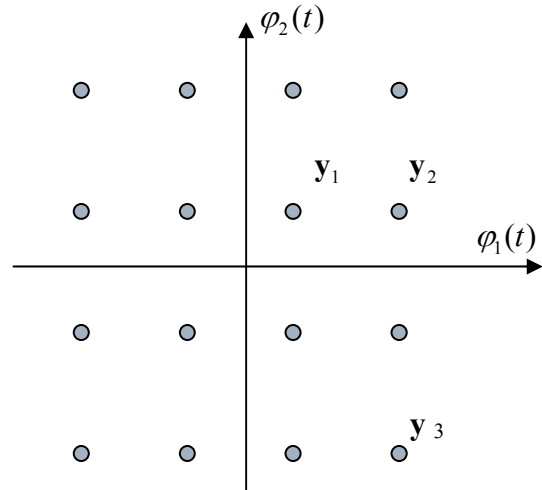
A QPSK-hoz hasonlóan a QAM egy jelentős többszintű moduláció. A névben a Q a kvadratúrát rövidíti, és hasonlóan az előzőhöz, itt is két egymásra merőleges jel, egy szinusz és egy koszinusz a főszereplő. Ezeknek a jeleknek az amplitúdóját változtatva, majd összeadva állítható elő a modulált jel.

Szemben az eddigiekkel, most rögzítsük le a jeltér bázisfüggvényeit:

$$\varphi_1(t) = \sqrt{2/\tau_0} \sin\left(2\pi/\tau_0 \cdot t\right); \quad \varphi_2(t) = \sqrt{2/\tau_0} \cos\left(2\pi/\tau_0 \cdot t\right), \quad 0 \leq t < \tau_0.$$

Könnyű észrevenni, hogy ezek megegyeznek a QPSK esetén használt bázisfüggvényekkel.

Nézzük például a 16QAM esetét, ami azt jelenti, hogy 16 különböző csatorna-szimbólumot használunk, így mindegyik maximum 4 bit információ továbbítására lehet képes, vagy másképpen fogalmazva, négy-négy bináris szimbólum képviselőre alkalmas. A mellékelt ábrán bemutattuk a jelvektorok egy lehetséges (és szokásos) elrendezését a jeltérben. Példaként megadjuk az ábrán jelölt vektorokat és az általuk leírt elemi jeleket:



$$\mathbf{y}_1 = (A\sqrt{\tau_0/2}, A\sqrt{\tau_0/2}),$$

$$\mathbf{y}_2 = (3A\sqrt{\tau_0/2}, A\sqrt{\tau_0/2}),$$

$$\mathbf{y}_3 = (3A\sqrt{\tau_0/2}, -3A\sqrt{\tau_0/2}),$$

illetve:

$$y_t^{(1)} = A \sin\left(2\pi/\tau_0 \cdot t\right) + A \cos\left(2\pi/\tau_0 \cdot t\right),$$

$$y_t^{(2)} = 3A \sin\left(2\pi/\tau_0 \cdot t\right) + A \cos\left(2\pi/\tau_0 \cdot t\right), \quad 0 \leq t < \tau_0$$

$$y_t^{(3)} = 3A \sin\left(2\pi/\tau_0 \cdot t\right) - 3A \cos\left(2\pi/\tau_0 \cdot t\right).$$

Ortogonalis jelek

Az eddigiekben már többször szerepelt, hogy a jeltér bázisfüggvényei ortogonalisak, sőt egységnyi energiájúra normáltak, most viszont olyan elemi jeleket mutatunk, amelyek egymásra is ortogonalisak lehetnek, ami azt jelenti, hogy a skalár-szorzatuk nulla.

Az FSK

Bináris esetben az elemi jeleket két különböző frekvenciájú szinuszos jel képviseli:

$$y_t^{(1)} = A \cdot \cos(\omega_1 \cdot t + \phi_1)$$

$$y_t^{(2)} = A \cdot \cos(\omega_2 \cdot t + \phi_2); \quad t \in [0, \tau_0).$$

A nem nulla kezdőfázisok nem bírnak jelentőséggel, így legyen $\phi_1 = \phi_2 = 0$, és az egyszerűség kedvéért $2\omega_1 \cdot \tau_0 = k \cdot \pi$, ahol $k = 1, 2, \dots$, valamint legyen $\omega_2 = \omega_1 + \Delta\omega$!

A két elemi jel természetesen megkülönböztethető a frekvenciája alapján bármekkora (véges) frekvenciakülönbség esetén, de több szempontból előnyös, ha a jelek ortogonalisak, ami az alább egyenlet teljesülése esetén áll fenn:

$$0 = \int_0^{\tau_0} y_t^{(1)} \cdot y_t^{(2)} dt = A^2 \int_0^{\tau_0} \cos(\omega_1 \cdot t) \cdot \cos(\omega_1 \cdot t + \Delta\omega \cdot t) dt .$$

Oldjuk meg az egyenletet $\Delta\omega$ -ra:

$$0 = \int_0^{\tau_0} \cos(2\omega_1 \cdot t + \Delta\omega \cdot t) dt + \int_0^{\tau_0} \cos(\Delta\omega \cdot t) dt = \frac{(2\omega_1 + \Delta\omega) \cdot \sin(\Delta\omega \cdot \tau_0) + \Delta\omega \cdot \sin[\tau_0(2\omega_1 + \Delta\omega)]}{\Delta\omega \cdot (2\omega_1 + \Delta\omega)} \stackrel{2\omega_1 \cdot \tau_0 = k \cdot \pi}{=} \\ = \frac{(k \cdot \pi / \tau_0 + \Delta\omega) \cdot \sin(\Delta\omega \cdot \tau_0) + \Delta\omega \cdot (-1)^k \cdot \sin[\Delta\omega \cdot \tau_0]}{\Delta\omega \cdot (k \cdot \pi / \tau_0 + \Delta\omega)}$$

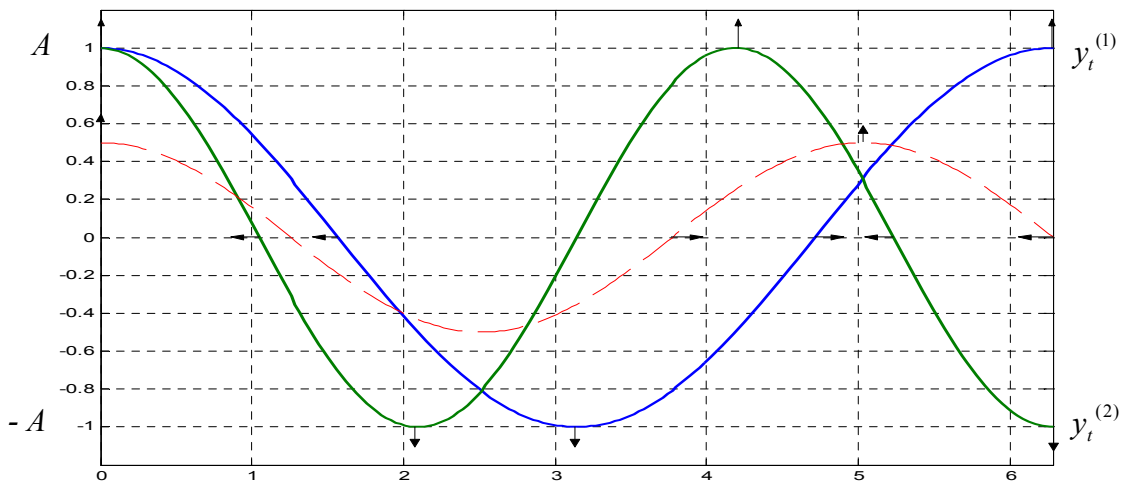
Mivel a kapott egyenlet jobboldalán szereplő kifejezésben a frekvenciák értelemszerűen nem lehetnek nullák, ezért megoldás csak a szinusz nulla értékeinél lehetséges, ami akkor következik be, ha:

$$\Delta\omega \cdot \tau_0 = \pi, \quad 2\pi, \quad 3\pi, \dots$$

Ezek közül a legkisebb értékű esetet *minimál fázisú* FSK-nak hívják, röviden MSK-val jelölik, és leggyakrabban ezt használják.

Ismerkedjünk meg kicsit részletesebben az MSK modulációval! Nézzük először az elemi jeleket, és az egyszerűség kedvéért legyen $k=1$ az alábbi összefüggésekben:

$$y_t^{(1)} = A \cdot \cos(k \cdot \frac{2\pi}{\tau_0} \cdot t) \\ y_t^{(2)} = A \cdot \cos(k \cdot \frac{2\pi}{\tau_0} \cdot t + \frac{\pi}{\tau_0} \cdot t) = A \cdot \cos[(2k+1) \frac{\pi}{\tau_0} \cdot t]; \quad t \in [0, \tau_0).$$

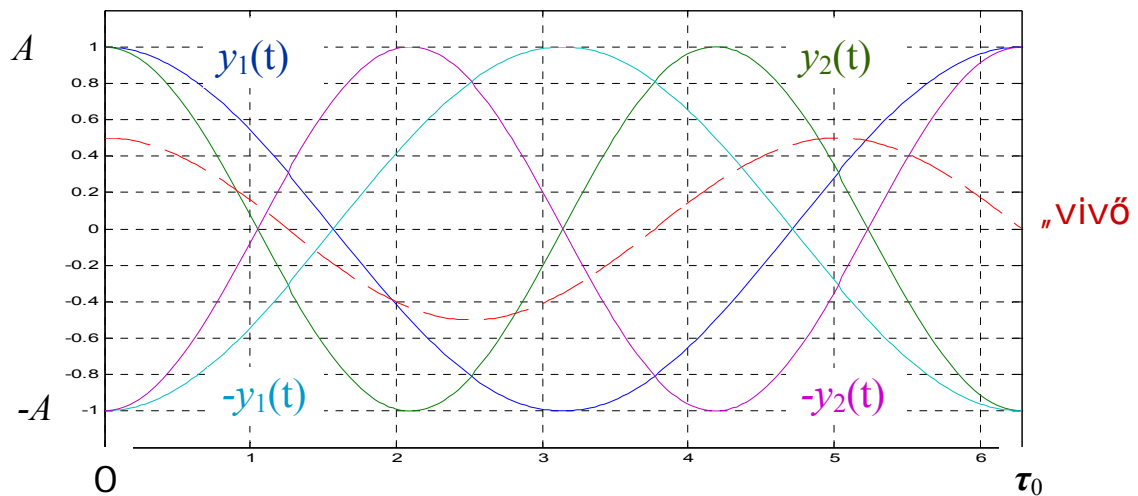


Az ábrán feltüntettük az MSK két elemi jelét az $\omega_1 = 2\pi/\tau_0$ esetén, valamint egy „elvi vivőt” (a szaggatott görbe), amelynek a frekvenciája a két, frekvenciában modulált elemi jel frekvenciájának a számtani közepe. Ennek a vivőnek a kezdőfázisát az 1-es elemi jelhez kötöttük, és ehhez a vivőhöz mérhetjük a modulált jel fázisát. Megállapíthatjuk, hogy az időrés kezdetén mindkét jel a vivővel azonos fázisú, majd az időrés végére az 1-es jel elmarad $\pi/2$ -öt, a 2-es jel viszont siet ugyanennyit. Ez persze nem meglepő, hiszen frekvenciamodulációról van szó. Az viszont már gondot jelent, hogy az időrés határain fázisugrás jöhet létre, attól függően, hogy milyen értékű a moduláló szimbólum. Ugyanis amennyiben 1-es jelet 1-es jel követ, akkor a fázis folytonos, de ha 1-es jelet 2-es jel követ, akkor π értékű fázisugrás jön létre a jelben az időrés határán, és szintén π értékű a fázisugrás, ha 2-es jelet bármelyik jel követi.

A fázisugrás vonalakat eredményez a spektrumban, ami kedvezőtlen, mert energiát emészt fel. Az ugrások azonban elkerülhetők egy kis ügyeskedéssel.

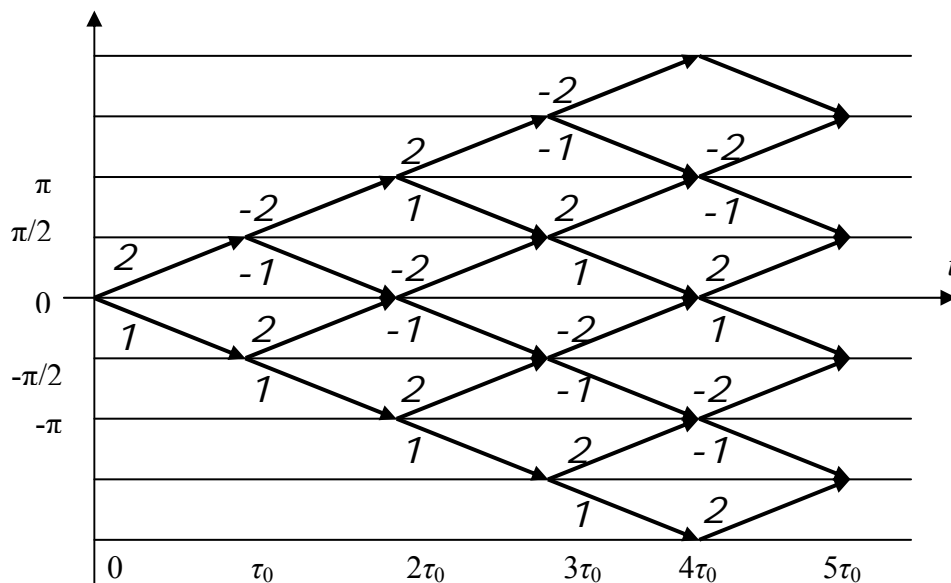
A folytonos fázisú MSK

Vezessünk be még két további elemi jelet, nevezetesen az eddigiek invertáltjait, azaz $-y_1^{(1)}$ -et, és $-y_1^{(2)}$ -öt, amint az alábbi ábra mutatja:



Könnyen ellenőrizhető, hogy megfelelően használva az elemi jeleket, fázisugrás nélküli modulált jelet lehet létrehozni. A demoduláció pedig csak kis mértékben lesz bonyolultabb, mivel két-két jelből kell egy-egy szimbólumra dönteni.

Mivel a frekvenciamoduláció esetén a fázisváltozás hordozza az információt, az FSK jel lényege fázis-trajektóriában is ábrázolható. Az alábbi ábra folytonos fázisú MSK fázisváltozását mutatja az idő függvényében:



A nullából indulva, 1-es szimbólum esetén $\pi/2$ fáziskésés lesz az időrés végén, 2-es szimbólum esetén ugyanekkora sietés. A diagram azt mutatja meg, hogy a folytonos fázis teljesüléséhez mikor kell a „pozitív” és mikor a „negatív” csatorna-szimbólumot használni. Például az egymást követő 2-esek esetén változtatni kell, vagy egy +2-est követően -1-nek kell jönni, stb. Az ábrából még az is kiolvasható, hogy a fázisváltozás az időrésben lineáris, ami persze nem szükségszerű, hiszen a digitális információ visszanyerése csak az időrésnek határain felvett értékekből történik. Spektrális szempontokból kedvezőbb lehet a fentitől eltérő fázis-trajektória is.

A folytonos fázisú MSK a jeltérben

Legyen például az egyik bázisfüggvény a következő:

$$\varphi_1(t) = \sqrt{\frac{2}{\tau_0}} \cdot \cos\left(\frac{2\pi}{\tau_0} \cdot t\right), \quad 0 \leq t < \tau_0.$$

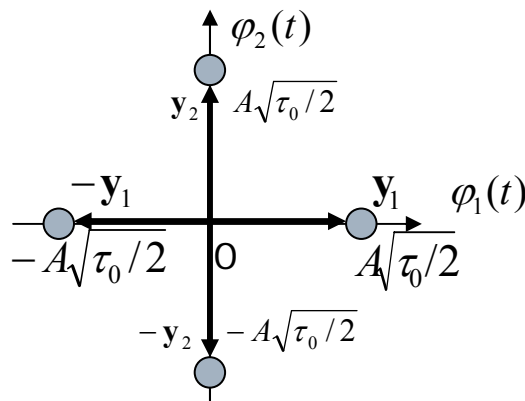
Az MSK-ra a másik bázisfüggvény az alábbi lesz, mivel $\Delta\omega = \pi/\tau_0$ -nál a jelek ortogonálisak:

$$\varphi_2(t) = \sqrt{\frac{2}{\tau_0}} \cdot \cos\left(\frac{3\pi}{\tau_0} \cdot t\right), \quad 0 \leq t < \tau_0.$$

A kétdimenziós jeltérben az elemi jelek a következők:

$$\mathbf{y}_1 = (A\sqrt{\tau_0}/2, 0),$$
$$\mathbf{y}_2 = (0, A\sqrt{\tau_0}/2).$$

Az így kapott két vektor még nem a folytonos fázisú MSK jelet írja le, hanem azt, amelynek fázisugrásai vannak a réshatárokon. Láttuk, hogy ennek megszüntetésére be kellett vezetni a fenti két elemi jel „tükörképeit” is, lesz tehát $-\mathbf{y}_1$ és $-\mathbf{y}_2$ is. Ezzel a folytonos fázisú MSK a jeltérben az alábbi vektorokkal jellemezhető:



Az ábrához két-három fontos megjegyzés tartozik: (i) annak ellenére, hogy négy vektor ábrázolja a jelet, a folytonos fázisú MSK egy bináris moduláció, de (ii) a fázis folytonossága érdekében cserélgetni kell az elemi jelet az előző időrés elemi jelétől függően, ami azt jelenti, hogy **nem** lesz **memóriátlan** a moduláció, illetve a csatorna, és (iii) bár a kép, amit az FSK jeltérbeli ábrája mutat, kísértetiesen hasonlít a QPSK-éra, de most különösen fontos megkülönböztetni a különböző fázisú jelek ú.n. konstellációját a jeltértől, hiszen itt az FSK-nál a fázis időben változó mennyiség, tehát az ábra bármely vektorának fázisa nem állandó.

Az FSK jel spektrális sűrűsége

Bár a spektrális sűrűség nem csekély erőfeszítések árán fontos esetekre kiszámítható, most megelégszünk néhány numerikus becslési eredmény rövid bemutatásával.

Ezeket az eredményeket úgy nyerhetjük, hogy álvéletlen szimbólumsorozatokkal modulált jelek FFT-jét átlagolva jó becslést kapunk a spektrális sűrűsége. Ezt az illusztrációt azonban felhasználjuk arra, hogy bemutassuk a folytonos fázisú FSK jel leírásának a korábbiaktól eltérő lehetőségét is. A jelenlegi leírás valójában nem az elemi jelekre koncentrálna, hanem ezen túlmenően a modulált jelet adjuk meg a következő kifejezéssel:

$$\eta_t = A \cdot \cos(\Omega \cdot t + \mu_t + \Phi),$$

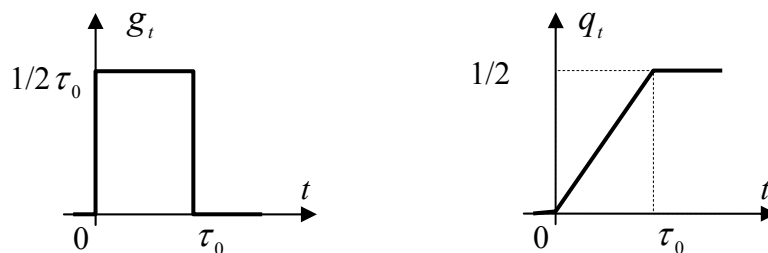
ahol Φ a véletlen kezdőfázis, amelyet most nullának fogunk tekinteni, és μ_t a fázismoduláció. Mivel a frekvenciaváltozást akarjuk arányossá tenni a moduláló jellel – lévén FSK a moduláció – a fázismoduláció a moduláló jel integráljával lesz arányos az alábbiak szerint:

$$\mu_t = m \cdot \pi \cdot \sum_{i=0}^{\infty} \xi_i \cdot q_{t-i\tau_0},$$

ahol

- $m := \frac{\Delta\omega \cdot \tau_0}{2 \cdot \pi}$ a modulációs tényező, a π egységekben mért fázislöket, amelyről megállapítottuk, hogy $\frac{1}{2}$ - es értékénél kapjuk az MSK-t (az MSK-nál π a fáziskülönbség a két elemi jel között az időrés végén, illetve fele ekkora a fáziseltérés a vivőhöz képest),
- ξ a moduláló – jelenleg bináris – szimbólumsorozat, és
- $q_t := \int_0^t g_s d\vartheta$ a fázisváltozás függvénye, amelyet a
- g_t ú.n. frekvencia-impulzussal szoktak megadni.

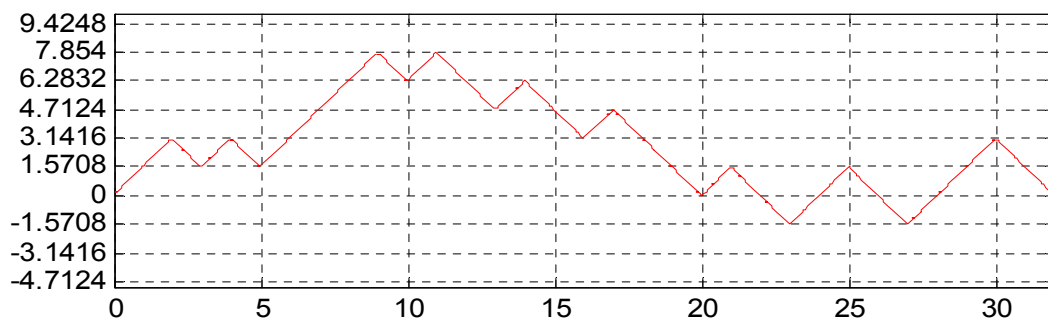
Az eddig vizsgált folytonos fázisú MSK jelet tekintve a fenti leírás elég bonyolultnak látszik, de összetettebb esetekre feltétlenül előnyös lesz. Példaként nézzük a folytonos fázisú MSK jelet, ahol a frekvencia-impulzus egy „közönséges” impulzus τ_0 hosszan, és akkora amplitúdóval, hogy a $\pi / 2$ értékű fázisváltozás létrejöjjön:



Legyen a ± 1 értékű bináris ξ sorozat például az alábbi:

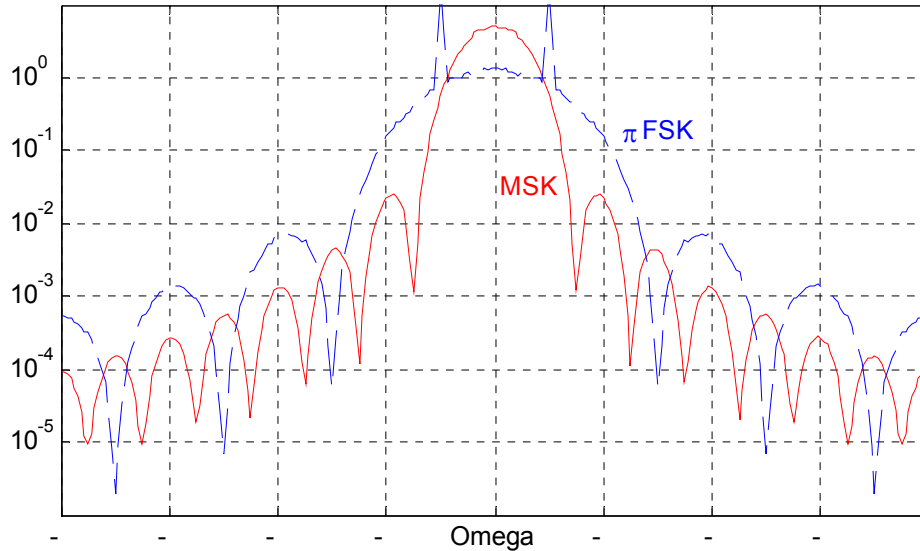
1 1 -1 1 -1 1 1 1 1 -1 1 -1 -1 1 -1 -1 1 -1 -1 1 1 -1 -1 1 1 -1 -1 1 1 1 -1 -1

a fázismoduláció így a következő lesz:

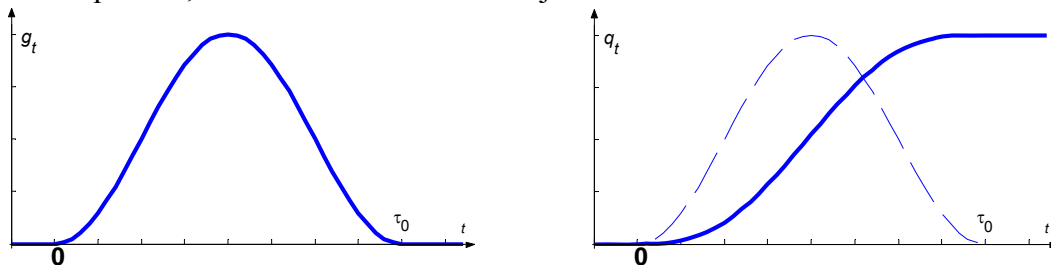


A görbe és a bináris sorozat pozíciói közel jól fedik egymást, így ellenőrizhető, hogy a +1 szimbólum $\pi / 2$ fázisnövekedést, a -1 szimbólum ugyanekkora csökkenést okoz. Magát a modulált jelet már nem ábrázoltuk, hanem áttérünk a spektrumra. Az alábbi ábrán láthatjuk a folytonos fázisú MSK jel spektrumát, valamint összehasonlításként feltüntetettük a π fázislöketet eredményező folytonos fázisú FSK spektrumát is (szaggatott vonal). Amennyiben a fázislöket $\pi / 2$ -nek nem páratlan számú többszöröse, akkor spektrumvonalak jelennek meg az elemi

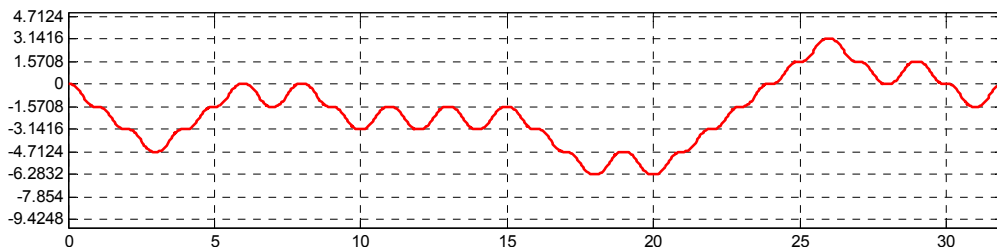
jelek frekvenciáinál. A folytonos fázisú MSK jel spektrumában ilyen nincs. Persze az is látható, hogy nagyobb fázislököt szélesebb spektrumot eredményez. Az ábrán az is látszik, hogy azonos teljesítményű jelekből számítva a spektrális sűrűséget, az MSK teljesítménye jobban koncentrálódik a vivő körül, magán a vivőn (Omega) több dB-vel nagyobb értékű, mint a viszonyításul választott π fázislökötű jel spektruma:



A sáv szélesség csökkentésének egyik lehetséges módja, ha nem ugráltatjuk a frekvenciát a két választott értéke között, hanem valamilyen simább átmenetet választunk. Talán a legkézenfekvőbb ötlet az lehet, hogy egy trigonometrikus függvényből alakítsuk ki a frekvencia-impulzust, amint az alábbi ábra mutatja:

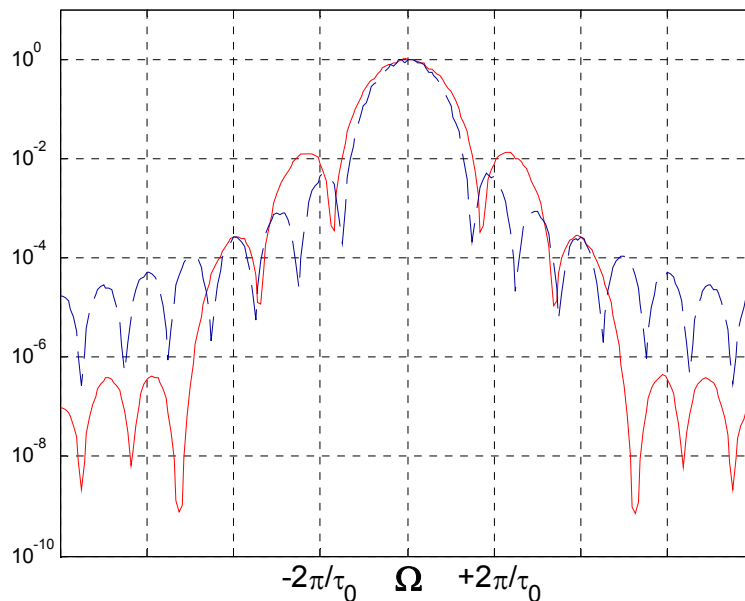


A jobboldali ábra görbéje a fázisváltozást írja le. Ismét bemutatunk egy véletlen sorozat által létrehozott fázismodulációs diagramot:



Az ábrán egy 32 bináris szimbólumból álló álvéletlen sorozat hatására létrejövő fázisváltozás függvényét láthatjuk, amely tehát sokkal „simábban” változik, mint az előbbi esetben és így várhatóan kisebb sáv szélességű modulált jelet eredményez.

A modulált jel becsült spektrális sűrűségét mutatjuk meg a következő ábrán, ahol viszonyításként az előbb meghatározott MSK spektrumot is feltüntettük szaggatott vonallal:



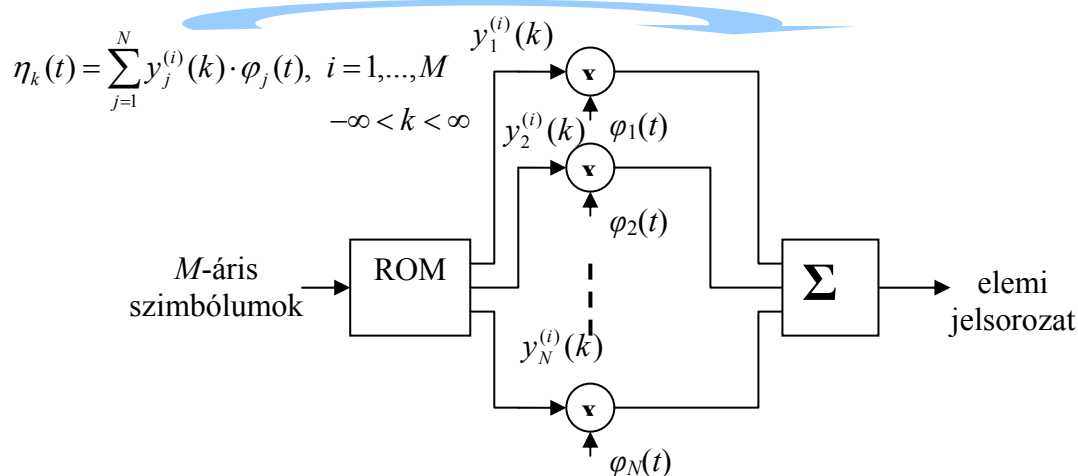
Az ábrán – a kényelem kedvéért – normálást végeztünk, azaz a vivőn felvett értékhez viszonyítottuk a spektrumot, továbbá bejelöltük a szimbólumtovábbítási sebességet is. Bár a sáv szélesség csökkentésének reményében változtattuk meg a modulációt – azt is mondhatjuk, hogy a változtatás valamiféle „szelídítés” volt – a kapott spektrum viszont „felemás” eredményt hozott: a közeli spektrális komponensek megnőttek, viszont a bitsebesség mintegy 2,5-szerese után mintegy 10 dB-vel kisebb spektrális komponenseket kaptunk.

További sáv szélesség-csökkentési lehetőségekre még a későbbiekben visszatérünk.

Összefoglaló

A jeltérbeli leírásra vonatkozó példák után rövid összefoglalásként bemutatunk egy szokásos interpretációt a modulációra és a demodulációra.

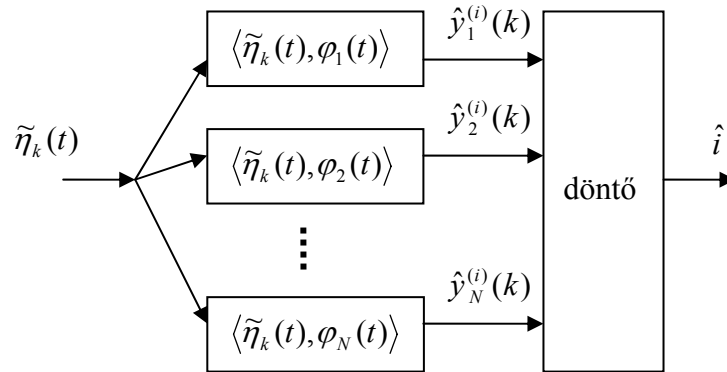
A korábban *szintézis egyenlet*-nek nevezett összefüggést tekinthetjük a digitális modulációnak az alábbi ábrán illusztrált módon:



Az $y_j^{(i)}(k)$ -val a szimbólumsorozat k -adik elemének megfelelő i jelvektor j -edik komponensét jelöltük.

Az *analízis egyenlet* pedig megfelel a demodulációnak. A k -adik időrésben érkező elemi jelről el kívánjuk dönteni, hogy melyik szimbólumot írja le, azaz i -re akarunk becslést adni:

$$y_j^{(i)}(k) = \int_0^{\tau_0} \eta_k(t) \cdot \varphi_j(t) dt; \quad i = 1, \dots, M, \quad j = 1, \dots, N; \quad \text{és} \quad -\infty < k < \infty$$



A döntéssel kapcsolatban meg kell jegyezni, hogy a modulált jelhez adódó zaj teszi a feladatot érdekessé, valamint korlátozott sávszélesség esetén a szomszédos időrészekben érkező jelek is befolyásolhatják az eredményt. Ezeket a kérdéseket a további fejezetekben vizsgáljuk. Jelenleg csak jeleztük a felső hullámmal, hogy a jel nem azonos a modulációval létrehozottal, de remélhetőleg hasonló hozzá, valamint a kalappal, hogy a jeltérbeli komponensek csak becslések.

Feladatok

1. Rajzolja le annak az FSK modemnek a négy elemi jelét, amelyik 1200 bit/sec átviteli sebességű, a kisebb frekvenciájú jel 600 Hz-es, és A amplitúdójú, továbbá folytonos fázisú MSK a moduláció! Adja meg a jelteret, a jel-vektorokat! Számítsa ki a jelvektorok közötti legkisebb távolságot! ($d = \sqrt{A^2 \cdot \tau_0}$)