

Debreceni Egyetem  
Informatika Kar

# Szakedolgozat

## Zajok és eltávolításuk

**Témavezető:** Dr. Papp Zoltán  
egyetemi adjunktus

**Készítette:** Gambár Péter István  
III. informatika tanár szakos,  
levelező tagozatos hallgató

Debrecen  
2009.

# Tartalomjegyzék

1. Bevezetés.....	4
1.1. A cím magyarázata.....	4
1.2. Témaválasztás indoklása .....	4
2. Zajok és modellezésük .....	5
2.1. Zajok fajtái és forrásai.....	5
2.1.1. Általános felosztás .....	5
2.1.2. Fehér zaj.....	7
2.1.3. Színezett zaj .....	7
2.1.4. Impulzív zaj.....	8
2.1.5. Tranziens zaj impulzus.....	10
2.1.6. Termikus zaj .....	11
2.1.7. Schottky-zaj.....	12
2.1.8. Elektromágneses zaj .....	13
2.1.9. Csatornatorzítások.....	14
2.1.10. Zajmodellezés .....	15
2.1.10.1. Hozzáadódó fehér normális eloszlású zajmodell (AWGN).....	16
2.1.10.2. Rejtett Markov modell.....	16
2.2. Zajmodellek .....	18
2.2.1. Valószínűségi modellek.....	18
2.2.1.1. Véletlenszerű jelek és sztochasztikus folyamatok.....	18
2.2.1.2. Valószínűségi modellek .....	19
2.2.1.3. Állandó és nem állandó véletlen folyamatok .....	23
2.2.1.4. Véletlen folyamatok várható értéke.....	24
2.2.2. A rejtett Markov modell.....	29
2.2.2.1. Nem stacionárius folyamatok statisztikai modellje .....	29
2.2.2.2. Rejtett Markov modellek.....	31
2.2.2.3. A rejtett Markov modell betanítása .....	34
3. Zajsűrítési módszerek.....	36
3.1. Zajzár .....	36
3.2. Analóg valós idejű zajsűrítők.....	36
3.2.1. Dolby zajsűrítő.....	36
3.2.2. Zajérvénytelenítés.....	38
3.3. Digitális jelek feldolgozása.....	39

---

3.3.1. Átalakítások .....	39
3.3.1.1. Analóg–digitális átalakítás .....	39
3.3.1.2. Digitális–analóg átalakítás .....	40
3.3.2. A Wiener szűrő.....	42
3.3.2.1. A legkisebb négyzetek becslése .....	42
3.3.3. Zajcsökkentés spektrális kivonással .....	44
3.3.3.1. A spektrális kivonás módszere.....	45
3.3.3.2. A spektrális kivonó megvalósítása.....	47
3.3.4. Az impulzív zaj eltávolítása.....	50
3.3.4.1. Impulzív zaj matematikai modellje .....	50
3.3.4.2. A medián szűrő.....	51
3.3.4.3. Impulzív zaj eltávolítása lineáris prediktív modell segítségével ...	53
4. Zajszűrés a gyakorlatban.....	58
4.1. Zenekari stúdiómunka .....	58
4.1.1. A felvétel .....	58
4.1.2. Az utómunka.....	59
4.2. Hagyományos hanglemez, hangkazetta archiválása CD-re.....	61
4.2.1. Hanglemez archiválása.....	61
4.2.1.1. A lejátszó rendszer .....	61
4.2.1.2. Várható zajok .....	63
4.2.1.3. Zajcsökkentési eljárás .....	64
4.2.2. Hangkazetta archiválása.....	65
4.2.2.1. A magnetofon és a hangszalag.....	65
4.2.2.2. Várható zajok .....	66
4.2.2.3. Zajcsökkentési eljárás .....	67
5. Összegzés .....	68
6. Irodalomjegyzék.....	70

## 1. Bevezetés

### *1.1. A cím magyarázata*

Általában hangfelvételek (függetlenül attól, hogy milyen eszközzel készülnek és milyen adathordozóra) és kommunikációs rendszerek esetében felütik fejüket az additív zajok. Ezen zajokat a hasznos jelhez szokták viszonyítani, és sajnos adott esetben a hasznos jel nagyságát is meghaladhatják, mely nagyban károsítja a jelek értelmezhetőségét. Több esetben azonban lehetőség nyílik az ilyen zajok mértékének, erősségének csökkentésére, mind hagyományos, mind digitális rendszerben. A zajok eltávolításánál nagyon fontos, hogy a hasznos jel minél nagyobb hányada maradjon ép, értelmezhető, és lehetőleg torzításmentes. Zajforrástól függően a zajok csökkentése hagyományos úgynevezett analóg rendszerekben is megvalósítható, akár megfelelő alkatrészek és eszközök alkalmazásával is. Jelen dolgozatban a tárolt és szállított hangjelek zajainak felismerésével, fajtaival és azok csökkentésével foglalkozom.

### *1.2. Témaválasztás indoklása*

Több éve egy zenei általános iskolában tanítok, ahol esetenként szükség van a kórus munkájának megörökítésére nem csak képi, de hang formájában is. A hangfelvételek megfelelő minőségű elkészítése és tárolása érdekében a rögzítő technika és az előtte található átalakító rendszerek által termelt zajok eltüntetése adott esetben nagyon fontos. Célom, hogy a dolgozat megírásával iskolámban minél hatékonyabb munkát tudjak végezni ebben a tekintetben is.

Személyes érdeklődési körömbé tartozik a zenei hangfelvételek digitális feldolgozása, illetve az erre alkalmas berendezések és szoftverek megismerése is, tehát többek között ez is motivált a szakdolgozat témájának megválasztásánál.

## 2. Zajok és modellezésük

### 2.1. Zajok fajtái és forrásai

#### 2.1.1. Általános felosztás

A zaj definiálható olyan nem kívánt jelnek, mely megakadályozza (zavarja) egy másik jellel (hasznos jel) történő kommunikációt vagy a másik jel mérését. A zaj maga is egy információhordozó, a zajforrás szempontjából. Például egy autó motorjának zaja információval szolgál a motor állapotáról. A zajok igen sokfélék és változatosak lehetnek, kezdve a hangfrekvenciás akusztikus zajoktól a rádiófrekvenciás elektromágneses zajokig.

A másik említésre méltó probléma a hasznos jel kiindulási alakjától való nem kívánatos eltérése, torzulása. Ebbe a csoportba tartoznak a különféle visszhangok, felharmonikusok, és az elveszett jelerészetek.

A zajok és torzulások a kommunikációs és mérési rendszerek fő ellenségei, ezért e káros mellékhatások eltüntetése érdemel kiemelt figyelmet. A zajszűrést és torzítás kiküszöbölését főként a mobil kommunikációban, a beszéd felismerésben, orvosi jelfeldolgozásban, radar és hanglokátor rendszerekben alkalmazzák a köznap életben.

A zajok életünk minden területén, minden környezetben jelen vannak, különféle formában. Az zaj forrásától függően a következő nagy csoportok különíthetők el egymástól:

- a.) *akusztikus zaj*: leggyakrabban mozgásból, vibrációból (rezgésből), ütközésből származik, ezáltal ez a zajforma – eltérő mértékben ugyan – jelen van a mindennapi környezetünkben. Ezt a zajt legtöbbször haladó gépjárművek, légkondicionálók, számítógépek hűtőventillátorai, közlekedés, emberi beszélgetések, szél, eső, stb. okozzák.
- b.) *elektromágneses zaj*: minden frekvencián jelen van, különösen a rádiófrekvenciákon. Minden elektromos berendezés, például a rádió- és TV adóállomások, és vevőkészülékek állítanak elő jelentős elektromágneses zajt.
- c.) *elektrosztatikus zaj*: feszültség jelenléte okozza, áram nélkül vagy folyó árammal. Ennek a zajnak az egyik legjelentősebb forrása a fénycsővilágítás.

d.) *csatornatorzítás, visszhang, és csillapítás*: az átviteli csatorna nem ideális karakterisztikája következtében történik. A rádiófrekvenciás csatornák, melyeket a mobil kommunikációban használnak, – mikrohullámú frekvenciákon – különösen érzékenyek a terjedés karakterisztikájára és környezetére.

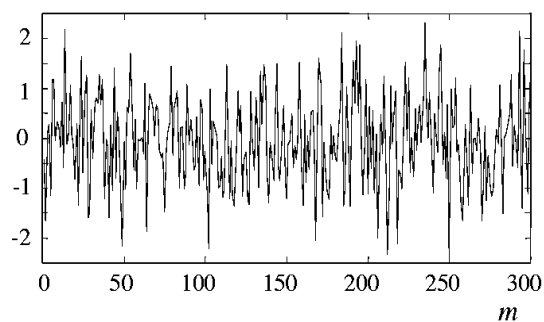
e.) *feldolgozási zaj*: az a zaj, amelyet a jel digitális–analog átalakítása eredményez, például a zaj digitálisan kódolt beszédben vagy hasonló jelekben, vagy az elveszett csomagok a digitális adatkommunikációs rendszerekben.

A zaj frekvenciájának vagy időkarakterisztikájának függvényében a következő kategóriákba csoportosítható:

- *Keskeny hullámsávú zaj*: olyan keskeny hullámsávon érkező zajhatás, mint az 50 Hz-es vagy 60 Hz-es brummogás (búgás) mely többnyire az elektromos tápegységekből, váltakozó áramú vezetékekből származik.
- *Fehér zaj*: tisztán véletlen zaj, melynek teljesen egyenes az energiaspektruma. A fehér zaj elméletileg az összes frekvenciát tartalmazza, azonos intenzitással.
- *Sávkorlátozott fehér zaj*: egyenes spektrumú zaj korlátozott sáv szélességgel, mely általában az alkalmazott eszköz korlátozott szélességű sávját, vagy az általunk használt sávot jelenti.
- *Szinezett zaj*: nemfehér zaj, vagy egyéb szélessávú zaj, melynek nem egyenes a spektruma, például: rózsaszín zaj, barna zaj, autoregresszív zaj.
- *Impulzusszerű zaj*: véletlenszerű rövid időtartamú impulzusokból áll, véletlenszerű amplitúdóval, hosszal és felbukkanással.
- *Tranziens zaj*: viszonylag hosszú időtartamú impulzusokból áll.

### 2.1.2. Fehér zaj

A fehér zaj definíció szerint egy összefüggéstelen jelekből álló zaj, mely minden frekvencián azonos energiájú. A zaj, melynek energiája minden frekvencián azonos  $\pm \infty$  tartományban, szükségképpen végtelen nagy energiájúnak kellene lennie, ennek következtében ez csak egy elméleti fogalom. Habár egy sávkorlátozott zaj,



1. ábra Fehér zaj

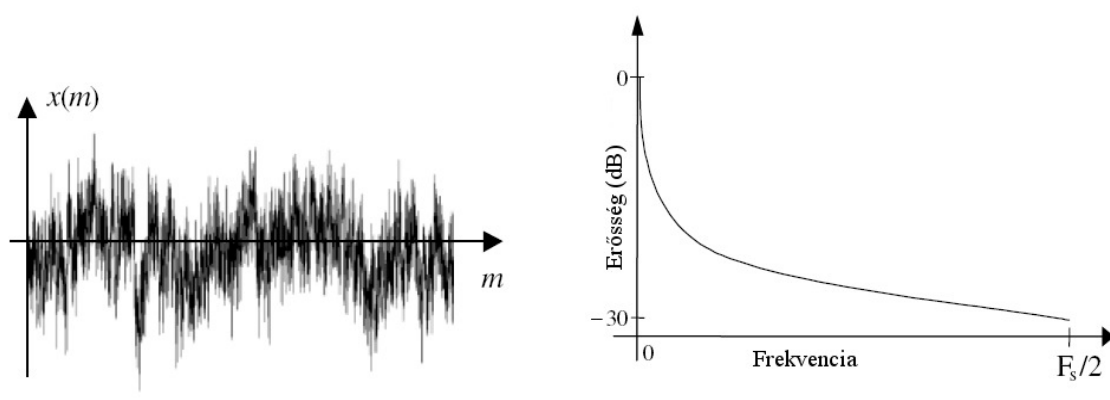
egyenes spektrummal – ha a zaj a teljes vizsgált sávot lefedi, amit a kommunikációs rendszer biztosít – a rendszer szempontjából minden tekintetben fehér zajnak számít. Például egy 10 kHz sávátvítellel rendelkező hangrendszer számára minden 10 kHz-nél szélesebb sávú egyenes spektrumú hallható zaj fehér zajnak

tűnik.

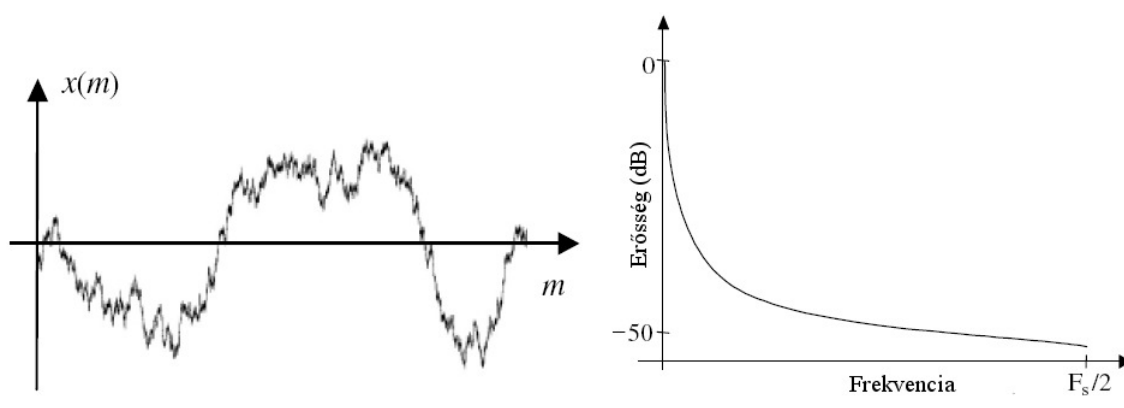
A tiszta fehér zaj egy elméleti fogalom, mivel végtelen energiájúnak kellene lennie a végtelen frekvenciatartomány lefedésére. Ráadásul a diszkrét időpontokban mintavételezett jelnek szükségképpen sávkorlátozottnak kell lennie, ahol a legnagyobb frekvenciaérték kisebb, mint a mintavételezési frekvencia fele. A sávkorlátozott fehér zaj egy sokkal gyakorlatibb fogalom, mely definíció szerint egy egyenes spektrumú zaj egy adott frekvenciasávban. Egy ilyen zaj energiája nem végtelen, ezért kiszámítható.

### 2.1.3. Színezett zaj

Bár a fehér zaj elmélete meglehetősen valós, és matematikailag kényelmes és használható közelítést ad több zajjal kapcsolatban, melyek felütik fejüket a telekommunikációs rendszerekben, sok egyéb zaj viszont nem fehér. A színezett zaj kifejezés bármely olyan szélessávú zajra utal, melynek spektruma nem a fehér zajéval egyező. Például a legtöbb hangfrekvenciás zaj, mint a mozgó autók zaja, a számítógépek ventilátorainak zaja, a fűrés zaja és emberi beszélgetés zaja, nemfehér, és túlnyomórészt alacsonyfrekvenciájú hangokat tartalmaz. Egy fehér zaj adott csatornán történő továbbításakor a zaj bizonyos frekvenciákon történő csillapodásával jellemezhető a „színezett zaj”, a csatorna spektrumának alakjától függően. A színezett zaj két klasszikus változata a rózsaszín (2. ábra) és a barna (3. ábra) zaj.



2. ábra Rózsaszín zaj és spektruma



3. ábra Barna zaj és spektruma

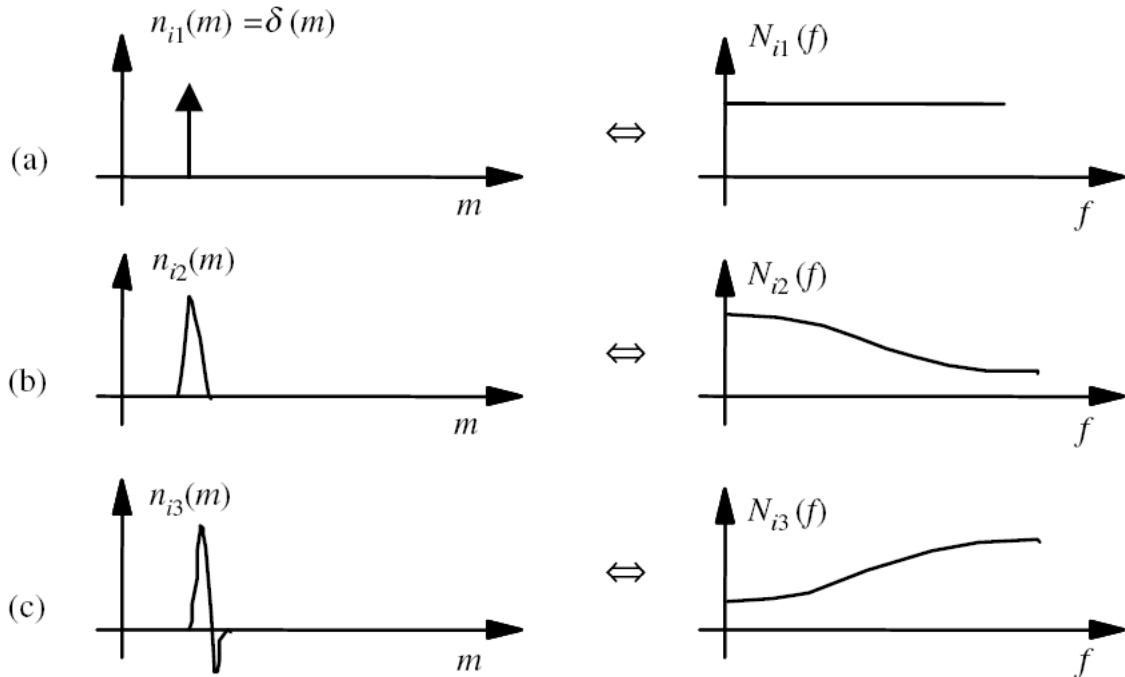
#### 2.1.4. Impulzív zaj

Az impulzív zaj rövid időtartamú „be/ki” impulzusokból áll. Ezek a zajok többféle forrásból származhatnak, mint például kapcsolási zaj, kedvezőtlen csatorna-környezet, kimaradás vagy felületi kopás a hanghordozón, számítógép-billentyűzet kattogása, és így tovább. A 4. ábra (a) része egy ideális impulzív zajt mutat, annak spektrumával. A kommunikációs rendszerekben előforduló impulzív zaj jellemző paramétere a hossza, mely normális esetben több mint egy minta hosszúságú. Például hangrendszerek esetében a rövid időtartamú, éles impulzusok lehetnek akár 3 milliszekundum (60 minta 20 kHz-es mintavételezési frekvencia esetében) hosszúak, melyeket impulzív zajnak tekinthetünk. A 4. ábra (b) és (c) része két példát mutat rövid időtartamú impulzusokra és azok spektrumára.

Egy kommunikációs rendszerben az impulzív zaj a térben és időben több pontból ered, azután a csatornán keresztül halad a vevő felé. A vett zaj időben szétszórta és a csatorna által alakított, és úgy kezelhető, mint a csatorna impulzus jelre adott reakciója. Általában a csatornák karakterisztikája lehet lineáris vagy nem lineáris,

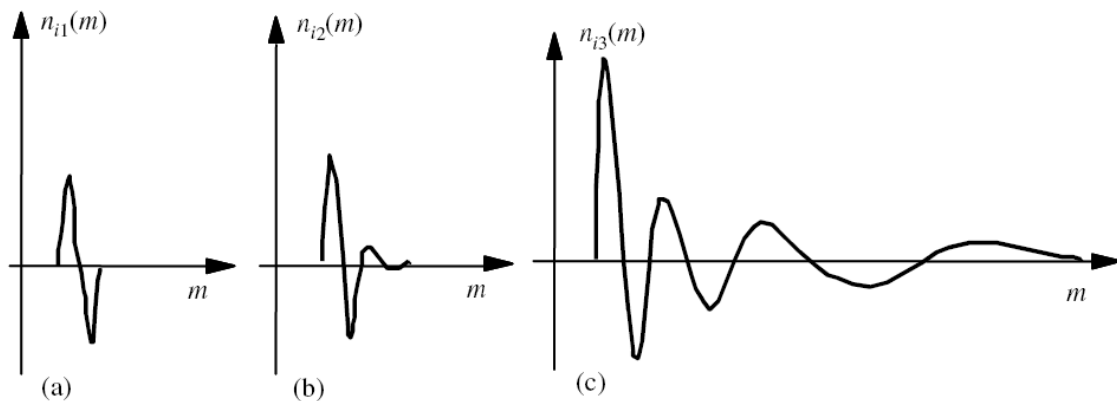


időben állandó, vagy időben változó. Azonkívül sok kommunikációs rendszer válaszul a nagy amplitúdójú impulzusokra nem lineáris karakterisztikát mutat.



4. ábra Impulzív zaj

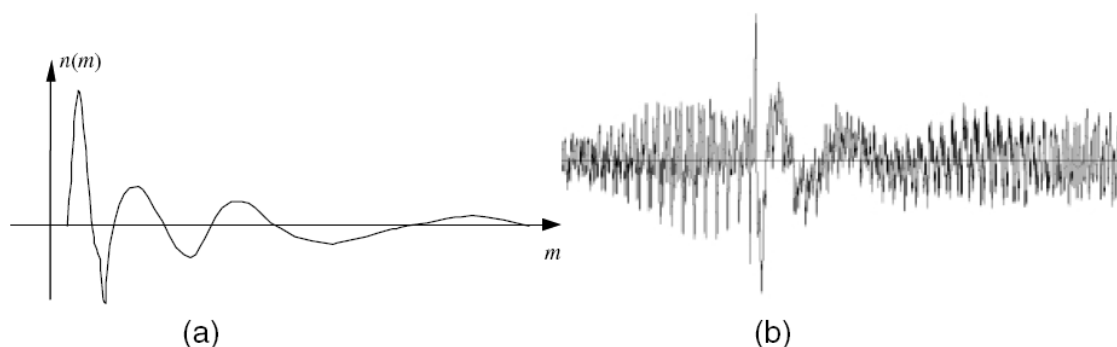
Az 5. ábra néhány példát mutat impulzív zajra, melyek tipikusan régi hanglemezeken fordulnak elő. Ebben az esetben a kommunikációs csatornának a lejátszó rendszert tekintjük, melyet időben állandónak feltételezhetünk. Az ábra három változatát mutatja a zaj-amplitúdóváltozásának az idő függvényében. Az 5. ábra (c) eleme mutatja, hogy egy nagy impulzussorozat csillapodó tranziens impulzussorozatot hoz létre. Ezek a példák a lejátszó rendszer nem lineáris karakterisztikáját jelzik.



5. ábra Példák impulzív zajra

### 2.1.5. Tranziens zaj impulzus

Az tranziens zaj gyakran áll egy viszonylag rövid éles impulzusból, melyet egy csillapodó viszonylag alacsony frekvenciájú oszcilláció követ, ahogy azt a 6. ábrán is láthatjuk. A kezdő impulzust gyakran valamilyen belső vagy külső interferencia okozza, míg az oszcillációt gyakran a csatorna rezonanciája, melyet a kezdeti impulzus gerjeszt, és tekinthető a csatorna válaszának a kezdeti impulzusra. Egy telekommunikációs rendszerben az impulzív zaj a térben és időben több pontból származik, azután a csatornán keresztül halad a vevő felé. A vett zaj időben szétszórt és a csatorna által torzított, alakított és úgy tekintendő, mint a csatorna egy impulzusra adott válaszreakciója. Az impulzív zaj a kezdő impulzussal és az azt követő oszcillációval jellemezhető, a csatorna szempontjából.



6. ábra Egy karc impulzus és zene egy hanglemezről (b), hanglemez karc impulzusának átlagolt profilja (a)

A 6. ábra jól szemlélteti a tranziens zaj impulzus alakját, tekintsük meg a karcolás impulzust<sup>1</sup> az ábrán, mely egy sérült hanglemezről származik. A karcolás impulzusok a lemeztű és a hozzá kapcsolt elektromechanikus lejátszó rendszer akusztikus válasza a lemezen található fizikai megszakítások jelenlétére. Mivel a karcolás hatására előálló tranziens zajok tulajdonképpen a lejátszó rendszer impulzus válasza, elvárható, hogy az adott rendszer különböző helyeken található impulzusokra hasonló karakterisztikát mutasson. Ahogyan a 6. (a) ábra is mutatja egy tipikus karc impulzus hullámalak gyakran két különböző részre tagolódik:

- egy bevezető nagy amplitúdójú lökésre, mely a lejátszó rendszer válasza a rögzített médiumon (lemezen) található karcolásra, melyet
- csillapodó hullámzások követnek, további torzítást okozva ezzel.

<sup>1</sup> Karc(olás) impulzus alatt egy hanglemezen, a használat során kialakult (többnyire a sávval valamilyen szöveget bezáró) karc (tekinthető pillanatnyi jelfolytonosság-hiánynak is) hatására létrejövő lejátszáskor jelentkező ugrásszerű lökést érthetünk, mely közvetlenül a mechanikus jelérzékelőre hat, és a mechanikai–elektronikai átalakítást követően végighalad a teljes elektronikus rendszeren is.

A kezdő impulzus általában viszonylag rövid, és a hossza a 1–5 ms tartományba esik, míg a hullámzó farok hosszabb időtartama elérheti akár az 50 ms-ot vagy többet is.

A 6. (b) ábrán látható csillapodó hullám (farok) frekvenciája is csökken az idő múlásával. Ez a viselkedés az elektromechanikus lejátszó-rendszer nem lineáris válaszmechanizmusa a lemezen található karcokra. Tapasztalat alapján több hanglemezről származó recsenés zaja jól meghatározható hullámalakot mutat, és viszonylag kevés tipikus sablonnal jellemezhető.

### 2.1.6. Termikus zaj

A termikus zaj – más néven Johnson-zaj (felfedezőjéről J. B. Johnsonról<sup>1</sup> elnevezve) – hővel stimulált részecskék rendezetlen mozgásából adódik. A termikus zaj elméletének alapjai a termodinamikában keresendők. A termodinamikában a szabad részecskéket hőmérsékletfüggő rendezetlen mozgás jellemzi, mint például a szabad gázmolekulák egy tartályban, vagy a szabad elektronok egy vezetőben. Bár ezen rendezetlen mozgások átlagértéke közel nulla, az átlag körüli hullámzások alkotják a termikus zajt. Például a gázmolekulák rendezetlen mozgásai és ütközései egy behatárolt térben rendezetlen hullámzást okoznak az átlagos nyomás körül. A hőmérséklet növekedésével a molekulák vagy más elemi részecskék kinetikai energiája növekszik, és ezzel a termikus zaj is.

Ehhez hasonlóan egy elektromos vezetőben nagy számú szabad elektron van jelen (az ionokkal együtt), melyek rendezetlenül rezegnek az egyensúlyi pozíciójuk körül, és útját állják a további elektronok mozgásának. Az elektronok rendezetlen szabad mozgásai spontán áramokat, vagy más szóval termikus zajt képeznek, melynek átlaga nulla, mivel feszültség hiányában minden elektron mozgása különböző irányú (nincs egyirányú áramlás). A környezet hőmérsékletének növekedésével a vezető hőmérséklete is növekszik, mely az elektronokat magasabb energiaszintre viszi, ezzel növelve a rendezetlen áramfolyásokat. Egy fémréteg ellenállásban a pillanatnyi feszültségek négyzetének átlagértéke a termikus zaj miatt a következő összefüggéssel adható meg:  $\overline{v^2} = 4kTRB$ , ahol  $k = 1,38 \cdot 10^{-23} \frac{\text{J}}{\text{°K}}$  a Boltzmann-állandó,  $T$  az abszolút hőmérséklet °K-ben,  $R$  az ellenállás ohmban, és  $B$  a sávszélesség. Az iménti egyenletből és az azt megelőző érvelésből egy áramkörü

<sup>1</sup> John Bertrand “Bert” Johnson (1887–1970): svéd születésű amerikai villamosmérnök és fizikus. Ő fejtette ki először részletesen a kábeleken közvetített információ véletlen interferenciájának alapjait. (Wikipédia – [http://en.wikipedia.org/wiki/John\\_B.\\_Johnson](http://en.wikipedia.org/wiki/John_B._Johnson))

lapba ültetett fémréteg ellenállás úgy viselkedik, mint egy termikus zajgenerátor, mely feszültségnégyzetének átlaga  $\overline{v^2}$ , és a belső ellenállása  $R$ . Az ily módon előálló termikus zajgenerátor teljesítménye számítható ki a következő összefüggéssel:

$$P_N = \overline{i^2} R = \left( \frac{v_{rms}}{2R} \right)^2 R = \frac{\overline{v^2}}{4R} = kTB[\text{W}], \text{ ahol } v_{rms}^1 \text{ az effektív feszültség értéke. A ter-}$$

mikus zaj spektrumsűrűsége megadható a  $P_N(f) = \frac{kT}{2} \left[ \frac{\text{W}}{\text{Hz}} \right]$  összefüggéssel. Ebből

az egyenletből kiderül, hogy a termikus zaj spektrumsűrűsége egyenes, azaz a termikus zaj egy fehér zajnak tekinthető (mely általában sávkorlátozott). Ezt az összefüggést különösen a magas rádiófrekvenciák ( $10^{13}$  Hz) esetében fontos.

### 2.1.7. Schottky-zaj

A Schottky-zaj<sup>2</sup> elmélete az elektroncsövek katódjának elektronkibocsátás-ingadozásából származik. Az különállónak tekintett elektronok áramlás során véletlenszerű időpillanatokban érkeznek, ezért az átlagérték körüli ingadozást okoznak a részecskefolyamban. Ezek az ingadozások a részecskefolyam sebességében Schottky-zajt okoznak. Egyéb példák Schottky-zajra a fotonok áramlása egy lézer fénysugárban, elektronok és lyukak áramlása valamint rekombinációja<sup>3</sup> félvezetőben, vagy a fotóelektronok árama, melyek a fotódiódákban emittálódnak. A véletlenszerűség jelenléte a részecske-kibocsátás vagy -érkezés mértékében magába foglalja, hogy a Schottky-zaj Poisson eloszlással modellezhető. A hosszú megfigyelési idejű kísérletek esetén az érkezések átlagérték körüli ingadozása a normális eloszláshoz közelít. Míg a termikus zaj a részecskék „nyugalmi állapotban” történő véletlenszerű mozgásából adódik, addig a Schottky-zaj egy (feszültséggel) irányított részecske-áramlásra értelmezhető.

<sup>1</sup>  $v_{rms}$  – root mean squared voltage: a feszültség effektív értéke

<sup>2</sup> Walter Hermann Schottky német fizikus, a film-rácsú vákuumcső (1915) és a tetróda (1919) feltalálója – munkáit a Siemensnél végezte. 1938-ban megfogalmazott egy elméletet, mely a Schottky-hatást jelezte, ma pedig a Schottky-diódákban használjuk. – [http://en.wikipedia.org/wiki/Walter\\_H.\\_Schottky](http://en.wikipedia.org/wiki/Walter_H._Schottky)

<sup>3</sup> „A félvezető nem csak szabad elektronjai útján vezeti az áramot (mint a fémek), hanem egy másik vezetési mechanizmus: a lyukvezetés útján is. Az eredetileg elektromosan semleges atom a negatív töltésű elektron távozása folytán pozitív töltésűvé válik, ezért a lyukat pozitív töltésűnek tekintjük. Előfordulhat, hogy egy szomszédos atom valamelyik vegyértékelektronja szintén elszabadul, és „beugrik” az előbb említett (pozitív töltésénél fogva az elektront vonzó) lyukba. Ezzel az előbbi lyuk (és az oda beugró szabad elektron is) megszűnik, hiszen a rácsszerkezet rendje az adott helyen helyreállt (ezt a folyamatot nevezik *rekombinációnak*), viszont a most elszabadult szabad elektron helyén új lyuk keletkezett. Ha valamelyik szomszédos atomból elszabaduló elektron most ebbe a lyukba ugrik be, tehát rekombinálódik, a lyuk a most elszabadult atom helyére »vándorol át«.” Forrás: Internet – [http://www.puskas.hu/r\\_tanfolyam/felvezetok.pdf](http://www.puskas.hu/r_tanfolyam/felvezetok.pdf), 2. oldal

Tekintsünk egy elektromos áramot, mint különálló elektromos töltéshordozók áramlását. Ha a töltéshordozók egymástól függetlenül mozognak, akkor az áram ingadozása megadható a  $I_{Noise}(rms) = \sqrt{2eI_{dc}B}$  összefüggéssel, ahol  $e = 1,6 \cdot 10^{-19}$  coulomb (egy elektron töltése), és  $B$  pedig a mérhető sávszélesség. Például egy  $I_{dc}$  1 amper „állandó” áram, és 1 MHz sávszélesség mellett, az effektív hullámzás 0,57 mikroamper lesz. Az iménti összefüggés feltételezi, hogy egymástól független töltéshordozók alkotják az áramot. Abban az eset mikor a töltéshordozóknak egy akadályon kell áthaladniuk, mint például egy félvezető dióda p–n átmenete, a töltéshordozók nem együtt mozognak; viszont nem igaz ez a fém vezetőkre, ahol a töltéshordozókat nagyfokú kölcsönhatás jellemzi.

### 2.1.8. Elektromágneses zaj

Virtuálisan minden elektromos eszköz, mely energiát állít elő, energiát fogyaszt vagy energiát szállít, elektromágneses zajok potenciális forrása. Az ilyen elektromágneses zaj interferenciának tekinthető más rendszerek szempontjából. Általában minél nagyobb a feszültség vagy az áram és minél közelebb helyezkedik el az áramkörhöz vagy elektromos eszközhöz, annál nagyobb lesz az áramkörben tapasztalható elektromágneses zaj mértéke. Az elektromágneses zajok általános forrásai a transzformátorok, rádió és televízió adóállomások, mobiltelefonok, mikrohullámú adóállomások, váltakozó áramú vezetékek, motorok és motorindítók, generátorok, relék, oszcillátorok, fénycsövek, és villámlások (zivatarok).

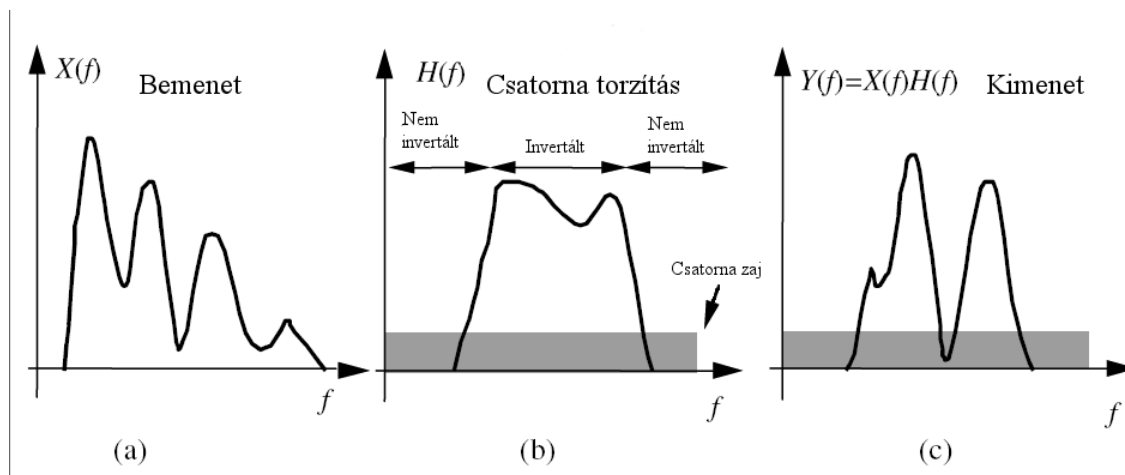
Az elektromágneses zajok két alapvető csoportba sorolhatók: elektrosztatikus zajok, és mágneses zajok. Ez két típus alapvetően különbözik egymástól, ezért más-más zajvédelem szükséges hozzájuk. Sajnos az imént felsorolt zajforrások legtöbbje a két típus kombinációját állítja elő, mely tovább bonyolítja a zajcsökkentési problémát.

Az elektrosztatikus tereket a feszültség jelenléte generálja, árammal, vagy anélkül. Az elektrosztatikus zajok egyik legáltalánosabb forrásai a fénycsövek. Mágneses terek keletkeznek elektromos áram hatására, vagy állandó mágnes jelenléte esetén. A motorok és a transzformátorok szolgálhatnak például az iménti felsorolásból, és hogy ne menjünk messzire a Föld mágneses tere is itt van környezetünkben. Ahhoz, hogy zajfeszültség indukálódjon egy vezetőben, a mágneses erővonalakat metszenie kell. A generátorok működési elve is ezen a jelenségen alapul. A váltakozó mágneses tér jelenléte – amilyen például az 50/60 Hz-es elektromos vezetékeket körülveszi – feszültséget indukál minden álló vezetőben, ahogy a mág-

neses tér kialakul, majd eltűnik (váltakozó mágneses tér). Ehhez hasonlóan, ha egy vezető mozog a Föld mágneses terében, zajfeszültség indukálódik benne, ahogy mozgás közben metszi a Föld mágneses erővonalait.

### 2.1.9. Csatornatorzítások

Ha egy jelet egy csatornán keresztül továbbítunk, a jel alakja változik, torzul a csatorna frekvenciaátvivő képességének és csillapítási paramétereinek megfelelően. A csatornatorzítás két fő megnyilvánulási formája a folytonossági torzítás és a fázistorzítás. A rádiókommunikációban létezik egy multi-út hatás is, mely szerint ha a jel több átjátszón megy keresztül, bizonyos esetekben különböző utakon haladva, ezzel a jel többféle változata érkezik a vevőbe különböző késleltetéssel és csillapítással. A csatornatorzítások csökkenthetők, vagy egészen egyszerűen szétválaszthatók a kommunikációs folyamatokban, ezért a csatorna modellezése és ki-

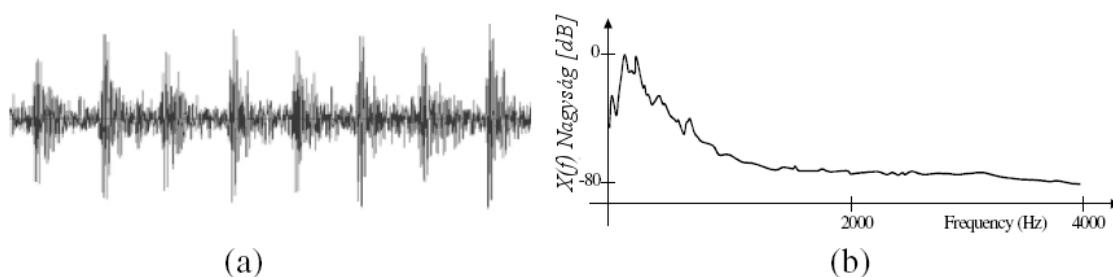


7. ábra Csatornatorzítás

egyenlítése nagyon jelentős összetevői egy modern digitális kommunikációs rendszer kialakításának. A csatorna kiegyenlítése különösen fontos a modern sejtes felépítésű kommunikációs rendszerekben mivel a csatorna karakterisztikák és terjedési paramétereik különbsége igen változatos, sokkal változatosabb, mint a vezetékes rendszerekben. A 7. ábra egy csatorna sávátviteli jellemzőit mutatja, mely két nem megfordítható és egy megfordítható régióra oszlik. A megfordítható csatornarégióban a jelek erősen torzultak elvesztek a csatornazaj miatt. A nem megfordítható csatornarégióban a jelek torzultak ugyan, visszanyerhetők. Ez a példa remekül mutatja, hogy a csatornafordító szűrőket nagy gondossággal kell elkészíteni, a nem kívánt eredmények elkerülése érdekében, mint például a zajerősítés az alacsony jel-zaj arányú frekvenciákon.

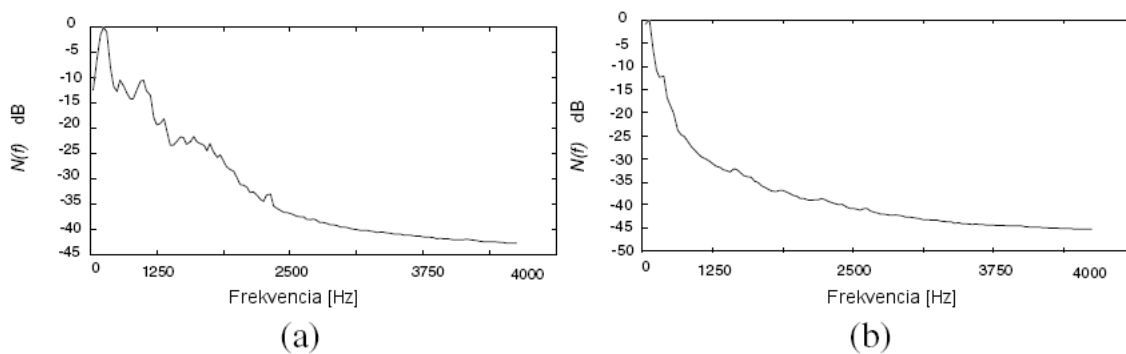
## 2.1.10. Zajmodellezés

A cél a modellezés esetén, hogy a jelben vagy zajban jelalakokat és mintákat jellemezzünk. A zaj pontos modellezéséhez egy olyan struktúrára van szükségünk, mely mind időben mind spektrumában jellemzi a zajt. A pontos zajstatisztika modellezéshez a kulcs a zajos jel megfelelő osztályozása és a hasznos jel erősítése. Ezen látszólag egyszerű jel/zaj osztályozási munka döntően függ a rendelkezésre álló hasznos jel mennyiségétől, erősségétől és a zajmodellektől, és ezen modellek alkalmazásától.



9. ábra A fúrózaj (a) idő-hullámformája (b) frekvenciaspektruma

A leghasználhatóbb szinte nélkülözhetetlen módszer az összes közül a zaj felépítésében való éleslátásra a Fourier transzformáció a zaj frekvenciaanalizálására. A 8. ábra egy elektromos fúró zaját mutatja, és ahogy az várható a zaj periodikus. Az ábra (b) része mutatja a fúró zajának spektrumát, melyen látható, hogy a zaj energiájának jelentős része a spektrum alacsonyabb frekvenciáira koncentrálódik. Tényszerűen igaz, hogy a legtöbb hangfrekvenciás jelnek és zajnak van túlnyomórészt alacsonyfrekvenciás a spektruma. Bár meg kell jegyezni, hogy a viszonylag kisebb energiájú magasabb frekvenciájú hangjelek fontos szerepet játszanak az élmény és a minőség átadásában. A 9. ábra (a) és (b) része példával szolgál az autók belsejéből származó zaj spektrumából (egy BMW és egy Volvo zajai láthatók). A



8. ábra Autók zajának spektuma (a) BMW 112 km/h, (b) Volvo 112 km/h

zajok nem statikusak, változatosak, és a következő forrásokból származhatnak:

- a.) kvázi periodikus zaj a motorból, és a jármű forgó alkatrészeiből,
- b.) zaj a kerekek és a talaj érintkezéséből és az út felületéből adódóan,
- c.) az autó körüli légáramlatok a légjáratokon, ablakokon, napfénytetőn keresztülhaladva,
- d.) az éppen előzőtt jármű zaja

Az autó zajainak jellemzői változnak a sebességgel, az út felületének állapotváltozásával, az időjárással és az autó környezetével.

A zajmodellezés legegyszerűbb módja, melyet gyakran használunk a jelenlegi gyakorlatban, hogy kiszámítjuk a zaj statisztikáját jelmentes időszakok alapján. Optimális esetben a Bayes féle jelfeldolgozási módszerekben egy készlet valószínűségi modell van előkészítve a jel- és zajfeldolgozásra. Ezek a modellek alapjául szolgálnak a jelek és zajok dekódolásának, zajos jelek felismerésének, erősítésének.

#### 2.1.10.1. Hozzáadódó fehér normális eloszlású zajmodell (AWGN<sup>1</sup>)

A kommunikációelméletben a gyakran feltételezzük, hogy a zaj hozzáadódó fehér normális eloszlást követő. Bár számos probléma esetében ez valós feltevés, és matematikailag kényelmes és használható megoldáshoz vezet, a gyakorlatban a zaj gyakran időben változó, kölcsönösen összefüggő, és nem normális eloszlást követő. Ez részben igaz az impulzív típusú zajokra, melyek nem állandóak, nem normális eloszlásúak, ezért nem modellezhetők az AWGN feltevés használatával. A nem állandó és nem normális eloszlású zajok modellezésére a Markov-féle sorozat alkalmas, stacionárius alfolyamatok leírásával.

#### 2.1.10.2. Rejtett Markov modell

A legtöbb zaj nem állandó, ez érvényes a zaj statisztikai paramétereire, úgy mint a zaj középértéke, szórásnégyzete, erőssége és spektruma, és ezek időbeli változása. A nem stacionárius zajok rejtett Markov modellek (HMM<sup>2</sup>) segítségével modellezhetők. Egy HMM lényegében stacionárius alfolyamatok véges sok állapotú Markov sorozata. Az állandó zaj modellezhető egy egyállapotú HMM segítségével. A nem állandó zaj modellezésére a többállapotú HMM alkalmas, amely képes a zaj időbeli változásának követésére, véges sok állandó állapottal. Nem normális eloszlású zajok esetében egy Gauss sűrűségi modell keverékét alkalmazhatjuk, a zaj terének modellezésére minden állapotban. Általában az állapotok száma modellen-

---

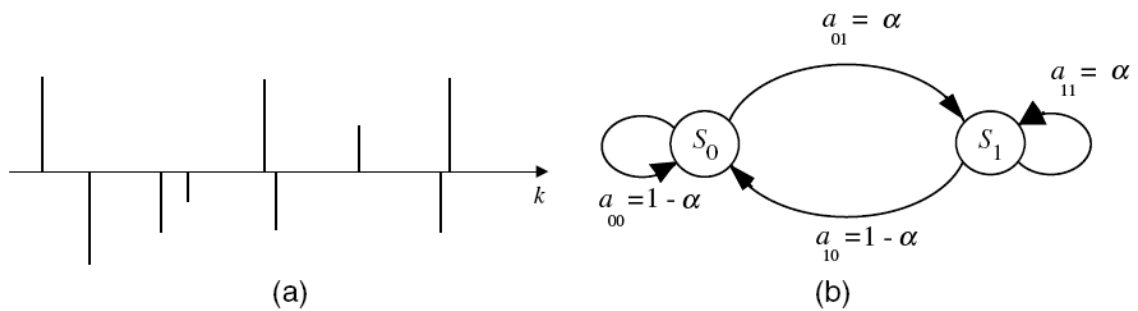
<sup>1</sup> AWGN: Additive White Gaussian Noise Model

<sup>2</sup> HMM: Hidden Markov Model



ként, és a keverékek száma állapotonként szükséges a zaj pontos modellezéséhez, mely a nem állandó zaj karakterisztikájától függ.

Egy példa a nem állandó zajra az impulzív zaj, melyet a 10. ábra (a) része szemléltet. Az ábra (b) részén az impulzív zaj szekvencia egy kétállapotú HMM modellje látható, ahol az  $S_0$  állapot jelképezi az „impulzus-ki” periódusokat az impulzusok között, és az  $S_1$  állapot jelképezi magát az impulzust. Azokban az esetekben ahol minden impulzus jól definiált átmeneti felépítéssel rendelkezik, előnyös lehet a többállapotú HMM magának az impulzusnak a modellezésére.



10. ábra (a) Impulzív zajszekvencia, (b) az impulzív zaj kétállapotú modellje

## 2.2. Zajmodellek

### 2.2.1. Valószínűségi modellek

A valószínűségi modellek alapja az információelmélet. Az információ maga is mennyiségileg a valószínűség logaritmusaként fejezhető ki. A valószínűségi modellek általában azzal foglalkoznak, hogy történések előfordulását megjósolják és jellemzésük véletlenszerű események alapján, különféle területeken, mint például telefonhívások számának megjóslása egy fővonalon a nap egy meghatározott időszakában, közúti forgalommodellezés, időjárás előrejelzés, pénzügyi adatmodellezés, előre jelezni a drogok hatását megadott adatok alapján az orvosi gyakorlatban. A jelfeldolgozásban a valószínűségi modelleket általában véletlenszerű jelek ingadozásának leírására alkalmazhatjuk, mint például mintafelismerés, jelkódolás és jelkiértékelés.

#### 2.2.1.1. Véletlenszerű jelek és sztochasztikus folyamatok

A jelek egyik legalapvetőbb jellemzőjük szempontjából két nagy csoportba osztályozhatók: determinisztikus (pontosan meghatározott) jelek és véletlenszerű jelek. Az időben véletlenszerű jelekre gyakran úgy hivatkozunk, hogy sztochasztikus jelek. Minden osztályban a jelek lehetnek folyamatosak és diszkrét az időben, és rendelkezhetnek folytonos értékű és diszkrét értékű amplitúdóval.

A determinisztikus jel definiálható úgy, hogy egyike az előre meghatározott görbéjű jeleknek időben és térben. A determinisztikus jelek pontos hullámmérete teljes mértékben leírható az idő függvényében, és a jel pontos értéke bármely időpillanatra kiszámítható a függvényi leírásból, és a korábbi értékek is kiszámíthatók ezen a módon. Például egy szinusz hullám  $x(t)$  modellezhető és pontosan kiszámítható mind egy másodrendű lineáris prediktív modellből, vagy a közismertebb  $x(t) = A \cdot \sin(2\pi ft + \phi)$  egyenlettel.

A véletlenszerű jeleknek nem kiszámítható hullámméretei vannak, ezért nem lehet olyan egyenletet felírni, mellyel kiszámítható lenne a jel pontos jövőbeli értéke a korábbi értékeiből. A legtöbb jel, mint például a beszéd, vagy zaj legalább részben véletlenszerű. A véletlenszerűség elmélete az információ és a zaj elméletével teljesen összekapcsolódott. Valóban, a véletlenszerű jelek feldolgozásakor a munka nagy részét jelenti az információ kinyerése a zajosan vett jelből. Ha egy jel alkalmas információhordozásra, akkor valamilyen fokú véletlenszerűségnek kell benne lennie: a kiszámítható jelek nem hordoznak információt. Ezért a jel véletlenszerű

része lehet a jel információtartalma, vagy zaj, vagy a kettő keveréke, tehát információ és zaj. Bár a véletlenszerű jelek nem teljesen kiszámíthatók, mégis gyakran mutatnak egy sor olyan statisztikailag jellemző értéket, mint például a maximum, a minimum, a középérték, a felező, a szórásnégyzet, és a teljesítményspektrum. A véletlenszerű jel leírható a statisztikája szempontjából, és a legteljesebben a valószínűségi modell szempontjából, melyből az összes statisztikája számítható.

#### 2.2.1.1.1. Sztochasztikus folyamatok

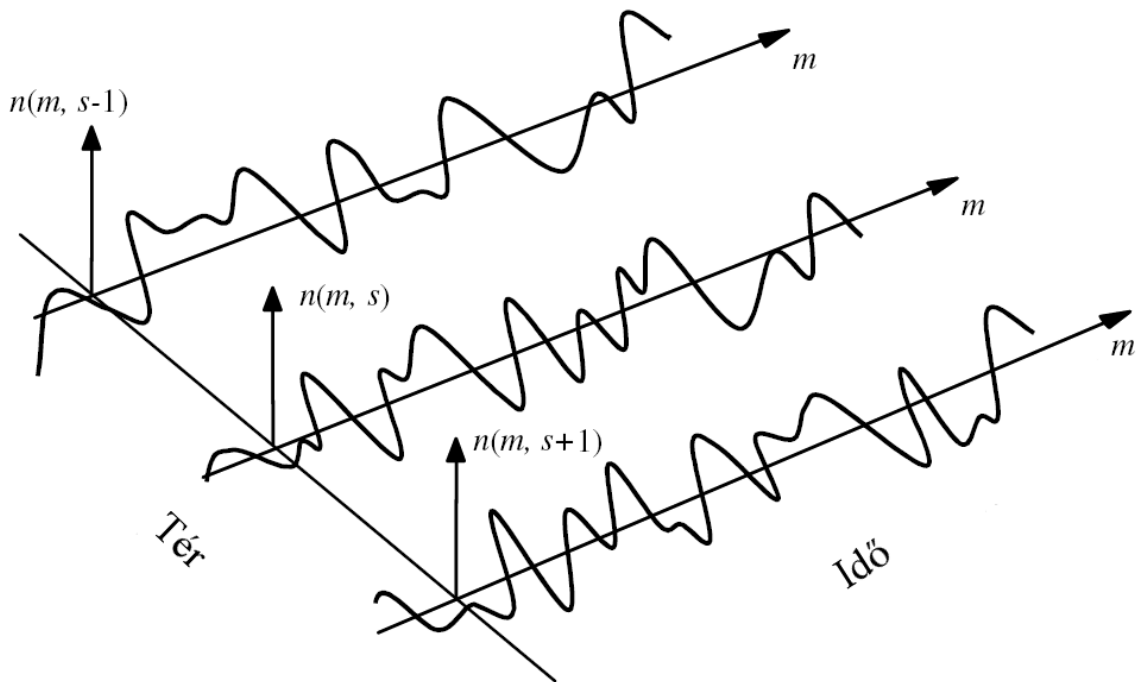
A „sztochasztikus eljárás” szakkifejezést általában a véletlenszerű folyamatok leírására használjuk, melyek folyamatos jeleket állítanak elő, mint például az emberi beszéd, vagy zaj. A jelfeldolgozási terminológiában egy sztochasztikus eljárás egy valószínűségi modell, a véletlenszerű jelek egy osztályára, például Gauss módszer, Markov módszer, Poisson módszer. Egy klasszikus példa a sztochasztikus eljárásra a részecskék úgynevezett Brown mozgása folyadékokban. A részecskék a folyadék belsejében véletlenszerűen mozognak a folyadék részecskék „bombázása” következtében. Minden részecske véletlenszerű mozgása a sztochasztikus módszer egy egyszerű elképzelése. Minden részecske folyadékban történő mozgása segíti a megértést, vagy a módszer különféle megvalósításainak terét.

#### 2.2.1.1.2. Véletlen folyamatok együttese vagy tere

A véletlenszerű folyamatok összes elképzelésének gyűjteménye a módszerek együttese vagy tere néven ismert. Példaként tekintsünk egy véletlenszerű zajt egy telekommunikációs rendszerben ahogyan azt a 11. ábra is mutatja. A zaj minden telefonvonalon időben véletlenszerűen ingadozik, és jelölhető az  $n(m,s)$  jelöléssel, ahol  $m$  a diszkrét idő mérőszám, és  $s$  jelöli a sorindexet. A zajok gyűjteménye a különböző vonalakon a zaj együttesét (vagy terét) alkotják, melyet a  $N(m)=\{n(m,s)\}$  jelöléssel jelölünk, ahol  $n(m,s)$  az  $N(m)$  zaj realizációját jelöli az  $s$  vonalon.

#### 2.2.1.2. Valószínűségi modellek

A valószínűségi modellek a véletlenszerű jelek legteljesebb matematikai leírását adják. Egy  $m$  rögzített időpillanatban egy véletlenszerű jeltől gyűjtött minta  $\{x(m,s)\}$  egy valószínűségi változó, mely különböző értékeket vesz fel az  $s$  térben. Az alapvető különbség a véletlenszerű változó és a véletlenszerű folyamat között, hogy az utóbbi idősorozatot generál. Ezért a véletlenszerű változók modellezésére használt modellek használhatók a véletlenszerű folyamatokra is.



11. ábra A véletlen zaj három realizációja a zajok terében

A valószínűségi változó tere az összes érték gyűjteménye, vagy eredmény, amit a változó felvehet. A valószínűségi változó tere részekre osztható, néhány kritérium szerint, néhány altérre. Az altér jelértékek egy gyűjteménye, melyek egy közös tulajdonsággal rendelkeznek, mint egy csoportba összezsúfolt minta, vagy a minták gyűjteménye azok amplitúdójával az adott sávon belül. Minden alteret egy eseménynek hívunk, és az esemény bekövetkezésének valószínűsége  $A$ , és  $P(A)$  az a hányados, mely az  $A$  térből származó megfigyelések eredményeinek számát  $N_A$ -t

elosztjuk az összes megfigyelés számával:  $P(A) = \frac{N_A}{\sum_i N_i}$ . Az iménti egyenletből evi-

dens, hogy a kísérletben résztvevő összes lehetséges esemény valószínűségének összege egy.

#### 2.2.1.2.1. Valószínűségi változó eloszlása, eloszlásfüggvény

Legyen adott egy  $(\Omega, \mathbf{A}, P)$  valószínűségi mező, legyen értelmezve egy  $\xi: \Omega \rightarrow \mathbf{R}$  függvény, melyre teljesül, hogy  $\{\omega \mid \xi(\omega) = x\} \in \mathbf{A}$  minden  $x \in \mathbf{R}$  esetén. ( $\xi$  az elemi események halmazát képezi le  $\mathbf{R}$ -be, értékkészlete diszkrét értékekből áll). Ekkor  $\xi$ -t diszkrét valószínűségi változónak nevezzük.  $\xi$  értékkészlete  $\{x_1, x_2, \dots\}$  megszámlálható.  $\xi$  eloszlása:  $\{p_i = P(\xi = x_i) = P(\omega \mid \xi(\omega) = x_i) \mid i=1, 2, \dots\}$  valószínűségek halmaza, annyi elemű, ahány elemű az értékkészlet, elemei mutatják a diszkrét valószínűségi változó egyes értékeinek felvételi valószínűségét. Az értékkészlet elemei nemnegatív számok, melyek összege 1.

Két diszkrét valószínűségi változó ( $\xi$  és  $\eta$ ) *együttes eloszlásán* a következő eloszlást értjük (amennyiben  $\xi$  értékészlete  $\{x_1, x_2, \dots\}$  eloszlása  $\{P(\xi=x_k) \ (k=1, 2, \dots)\}$ ,  $\eta$  értékészlete  $\{y_1, y_2, \dots\}$  eloszlása  $\{P(\eta=y_j) \ (j=1, 2, \dots)\}$ ):  $\{P(\xi=x_k \text{ és } \eta=y_j) \ (k,j=1, 2, \dots)\}$ .

Az együttes eloszlás elemei nemnegatívak, az elemek összege 1. Az együttes eloszlásban szereplő események teljes eseményrendszert alkotnak, ezért az előző állítás igaz.

A peremeloszlás két diszkrét valószínűségi változó együttes eloszlásából meghatározható az alkotó diszkrét valószínűségi változók eloszlása. Valamelyik diszkrét valószínűségi változó eloszlása szerint összegezve az együttes eloszlást, a másik diszkrét valószínűségi változó eloszlásához jutunk.

Diszkrét valószínűségi *változók függetlenek*, ha minden  $j = 1, 2, \dots$  és minden  $k = 1, 2, \dots$  esetén  $P(\xi=x_k \text{ és } \eta=y_j)=P(\xi=x_k)P(\eta=y_j)$ , azaz együttes eloszlásuk minden eleme szétbomlik az eloszlások megfelelő elemeinek szorzatára. Két diszkrét valószínűségi változó eloszlása pontosan akkor határozza meg együttes eloszlásukat, ha függetlenek. A konstans értékű valószínűségi változó minden más diszkrét valószínűségi változótól független.

Adott valószínűségi mezőben lévő  $B_i \ (i=1, 2, \dots)$  teljes eseményrendszer esetén ha ezek között nincs 0 valószínűségű esemény, akkor tetszőleges nem 0 valószínűségű  $A$  eseményre  $P(A | B) = \frac{P(A | B_i) \cdot P(B_i)}{\sum_i (P(A | B_i) \cdot P(B_i))}$ . Ez Bayes tétele.

Egy  $\xi$  valószínűségi változó *eloszlásfüggvénye*:  $F: \mathbf{R} \rightarrow [0, 1]$ ,  $F(x)=P(\xi < x)$ . Valószínűségi változó eloszlásfüggvénye balról folytonos, monoton nemcsökkenő,  $-\infty$ -ben 0-hoz tart,  $+\infty$ -ben 1-hez tart. Bármely, a valószínűségi változó eloszlásfüggvényeire vonatkozó fenti feltételeket kielégítő függvényhez létezik olyan valószínűségi változó, melynek eloszlásfüggvénye az adott függvény. A diszkrét valószínűségi változónak is létezik eloszlásfüggvénye, ez „lépcsős” alakú.  $\xi$  valószínűségi változó esetén  $P(a \leq \xi < b)=P(\xi < b)-P(\xi < a)=F(b)-F(a)$ , ha  $a < b$ .  $\xi$  és  $\eta$  *valószínűségi változók együttes eloszlásfüggvénye*  $F(x,y)=P(\xi < x \text{ és } \eta < y) \ (x,y \in \mathbf{R})$ .  $\xi$  és  $\eta$  *valószínűségi változók függetlenek*, ha együttes eloszlásfüggvényük megegyezik eloszlásfüggvényeik szorzatával. Valószínűségi változók együttes eloszlásfüggvénye mindkét változójában monoton nemcsökkenő, balról folytonos,  $\lim_{x \rightarrow -\infty} F(x,y)=\lim_{y \rightarrow -\infty} F(x,y)=0$ ,  $\lim_{y \rightarrow \infty, x \rightarrow \infty} F(x,y)=1$ ; ha  $a_1 < b_1$  és  $a_2 < b_2$ , akkor  $F(b_1,b_2)-F(a_1,b_2)-F(b_1,a_2)-F(a_1,a_2) \geq 0$ . Bármely, a valószínűségi változók együttes eloszlásfüggvényére vonatkozó ezen feltételeket kielégítő függvényhez létezik két olyan valószínűségi változó, melyek-

nek együttes eloszlásfüggvénye az adott függvény. Az utolsó feltétel  $P(\xi \in [a_1, b_1] \text{ és } \eta \in [a_2, b_2]) \geq 0$ -t jelenti.

### 2.2.1.2.2. Sűrűségfüggvény

Egy adott eseménytéren értelmezett  $\xi: \Omega \rightarrow \mathbf{R}$  függvény *valószínűségi változó*, ha  $\{\omega \mid \xi(\omega) < x\} \in \mathbf{A}$  bármely  $x \in \mathbf{R}$  esetén. A diszkrét valószínűségi változó teljesíti az iménti definíció követelményeit. A definíció a diszkrét valószínűségi változók általánosítása: folytonos esetben  $P(\xi=x)=0$  minden  $x \in \mathbf{R}$  esetén. Valószínűségi változót (ha nem hangsúlyozottan diszkrét) a definíció szerinti értelemben kell érteni.

Egy  $\xi$  valószínűségi változó *folytonos*, ha eloszlásfüggvénye abszolút folytonos. Folytonos valószínűségi változó esetén  $P(\xi=x)=0$  bármely  $x \in \mathbf{R}$  esetén.

Belátható, hogy  $0 \leq P(\xi=x) = \lim_{\Delta x \rightarrow 0} P(x \leq \xi < x + \Delta x) = \lim_{\Delta x \rightarrow 0} (F(x + \Delta x) - F(x)) \leq \lim_{\Delta x \rightarrow 0} ((F(x) + K\Delta x) - F(x)) = \lim_{\Delta x \rightarrow 0} (K\Delta x) = 0$ .

Bármely  $\xi$  valószínűségi változó esetén bármely  $x \in \mathbf{R}$ -re  $P(a \leq \xi \leq b) = F(b+0) - F(a)$ ,  $P(a < \xi < b) = F(b) - F(a-0)$ ,  $P(a < \xi \leq b) = F(b+0) - F(a-0)$  teljesül. Folytonos esetben ezek mind egyenlők  $P(a \leq \xi < b)$ -vel. Folytonos  $\xi$  valószínűségi változó *sűrűségfüggvénye*

$f: \mathbf{R} \rightarrow \mathbf{R}$ , ha  $\int_{-\infty}^x f(t) dt = F(x)$  teljesül minden  $x \in \mathbf{R}$  esetén. Folytonos valószínűségi

változó sűrűségfüggvényére és eloszlásfüggvényére teljesül az  $F'(x)=f(x)$  összefüggés. Diszkrét valószínűségi változónak nincs sűrűségfüggvénye.

A sűrűségfüggvény nemnegatív, az egész számegegyenesen vett integrálja 1. Bármely, a valószínűségi változó sűrűségfüggvényére vonatkozó, ezen feltételeket kielégítő függvényhez létezik olyan valószínűségi változó, melynek sűrűségfüggvénye az adott függvény. Diszkrét esetben hasonló feltétel a  $p_i \geq 0$ ,  $\sum p_i = 1$  tulajdonságok: az eloszlás játssza a sűrűségfüggvény szerepét is.

$\xi$  és  $\eta$  abszolút folytonos *valószínűségi változók együttes sűrűségfüggvénye*

$f(u, v)$ , ha rá teljesül az  $\int_{-\infty}^x \int_{-\infty}^y f(u, v) dv du = F_{\xi, \eta}(x, y)$  egyenlőség.

Az együttes sűrűségfüggvény megegyezik az együttes eloszlásfüggvény második parciális deriváltjával (először az első, majd a második változó szerint deriválva). Abszolút folytonos  $\xi$  és  $\eta$  valószínűségi változók esetén függetlenségük egyenértékű  $f_{\xi, \eta}(u, v) = f_{\xi}(u)f_{\eta}(v)$  teljesülésével. Az együttes sűrűségfüggvény nemnegatív, az egész valós számsíkon vett integrálja 1. Az együttes sűrűségfüggvényből meghatározható az alkotó valószínűségi változók sűrűségfüggvénye. Az együttes sűrűség-

függvényt az egyik változó szerint az egész valós számegyenesen integrálva a másik valószínűségi változó sűrűségfüggvényét kapjuk.

### 2.2.1.3. Állandó és nem állandó véletlen folyamatok

Bár a jel  $x(m)$ -mel jelölt amplitúdója  $m$  periódussal hullámzik, a jel által generált folyamat jellemzője lehet időben állandó vagy időben változó. Nem állandó folyamat például a beszéd, melynek hangossága és spektrális jellemzői, összetevői folyamatosan változnak, ahogyan a beszélő különféle hangokat képez. A folyamat állandónak tekinthető, ha a valószínűségi modellbeli jellemzői idő invariánsak, ellenkező esetben nem állandó – mindkettőt a 12. ábra szemlélteti. Az állandó tulajdonság magába foglalja, hogy a jel összes paramétere, úgy mint középérték, variancia, energiaspektrum összetétele, és a folyamat magasabb rendű momentumai idő invariánsak. A gyakorlatban az állandóságnak különféle fokai ismeretesek. Van olyan, hogy a statisztikák egy halmaza állandó, a többi pedig nem. Például lehet egy véletlen folyamatnak idő invariáns középértéke, de az idő függvényében változó energiája.



12. ábra Kvázistacionárius és nemstacionárius beszédsgemensek

### 2.2.1.3.1. Szigorú értelemben vett állandó folyamatok

Egy  $X(m)$  véletlenszerű folyamat szigorú értelemben véve stacionárius, ha minden megoszlása, és statisztikai paramétere idő invariáns. A szigorú értelemben vett állandóság magában foglalja, hogy az  $n$ -edik rendű eloszlása transláció invariáns minden  $n = 1, 2, \dots$ , vagyis:

$$P[x(m_1) \leq x_1, x(m_2) \leq x_2, \dots, x(m_n) \leq x_n] = P[x(m_1 + \tau) \leq x_1, x(m_2 + \tau) \leq x_2, \dots, x(m_n + \tau) \leq x_n]$$

Az egyenletből a szigorú értelemben vett állandó folyamat statisztikája beleértve a középértéket, a korrelációt és az energiaspektrumot, idő invariáns lesz, ezért:  $E[x(m)] = \mu_x$ ,  $E[x(m)x(m+k)] = r_{xx}(k)$  és  $E[|X(f, m)|^2] = E[|X(f)|^2] = P_{xx}(f)$ , ahol  $\mu_{xx}$ ,  $r_{xx}(m)$  és  $P_{xx}(f)$  rendre az  $x(m)$  jel középértéke, az autokorrelációja és az energiaspektruma, és  $X(f, m)$  jelöli  $x(m)$  frekvencia-idő spektrumát.

### 2.2.1.3.2. Nem szigorú értelemben vett állandó folyamatok

Míg a szigorú értelemben vett állandó (stacionárius) folyamatok esetében a minden statisztikának idő-invariánsnak kell lennie, a nem szigorú értelemben vett stacionárius folyamatokra nem vonatkozik az összes korlátozás. Egy ilyen folyamat középértéke és az autokorrelációs függvénye idő invariáns:  $E[x(m)] = \mu_x$ ,  $E[x(m)x(m+k)] = r_{xx}(k)$ . A két folyamat definíciójából kiderül, hogy minden szigorú értelemben vett stacionárius folyamat nem szigorú értelemben is az, viszont ennek a fordítottja nem igaz.

### 2.2.1.3.3. Nem stacionárius folyamatok

Egy véletlenszerű folyamat nem stacionárius, ha eloszlása vagy statisztikája időben változik. A legtöbb sztochasztikus folyamat, mint például a videojelek, audiojelek, gazdasági adatok, meteorológiai adatok, stb., nem stacionárius, mert egy-egy olyan rendszer generálja őket, melynek környezeti paraméterei időben változnak. Például a beszéd nem stacionárius, mert egy időben változó artikulációs rendszer állítja elő. A beszéd hangerőssége, frekvencia-összetétele időben változó, mely változás néha teljesen váratlan. Az időben változó folyamatok modellezhetők stacionárius véletlenszerű folyamatok egy kombinációjával.

### 2.2.1.4. Véletlen folyamatok várható értéke

A várható érték központi szerepet játszik a jelek modellezésében és feldolgozásában. Továbbá a véletlenszerű folyamatok valószínűségi modelljei általában a várható érték függvényeként kerülnek kifejezésre. Például a normális eloszlás el-



oszlásfüggvénye a folyamat középértékét és kovarianciáját tekintve egy exponenciális függvényként definiálható, a Poisson eloszlás eloszlásfüggvénye pedig a folyamat középértékével definiálható. A jelfeldolgozási alkalmazásokban mindig megvan a jel megfelelő statisztikai modellje – például a normális eloszlás eloszlásfüggvénye –, és a modell kiegészítésére szükség van a várható paraméterek értékére. Azonkívül sok jelfeldolgozási algoritmusban, mint például a spektrális kivonásban a zajcsökkentésre, vagy a lineáris előrejelzésben, lényegében amire szükségünk van az a jel középértékének vagy a korrelációs függvényének kiszámítása. Egy véletlenszerű folyamat  $h(X(m_1), X(m_2), \dots, X(m_M))$  függvényének várható értékének kiszámítása a következőképpen definiálható:

$$E[h(X(m_1), \dots, X(m_M))] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} h(x_1, \dots, x_M) f_{X(m_1) \dots X(m_M)}(x_1, \dots, x_M) dx_1 \dots dx_M$$

A legfontosabb és széles körben használt várható értékek a középérték, a korreláció, a kovariancia és a energiaspektrum.

#### 2.2.1.4.1. A középérték

A egy jelsorozat középértéke nagyon fontos szerepet játszik a jelfeldolgozásban és paraméterkiszámításban a zajos jelek esetében. Például egy zajos jel esetében az optimális lineáris kiszámítási módszer az interpoláció a középérték és a megfigyelt jel zajjal terhelt értéke között. Egy  $[X(m_1), \dots, X(m_M)]$  véletlenszerű vektor középértéke alatt értjük a teljes jelfolyam alatt mutatott átlagértéket, mely a következőképpen definiálható:

$$E[X(m_1), \dots, X(m_M)] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (x_1, \dots, x_M) f_{X(m_1) \dots X(m_M)}(x_1, \dots, x_M) dx_1 \dots dx_M$$

#### 2.2.1.4.2. Autokorreláció

A korrelációs függvény és annak Fourier transzformációja, az energiaspektrum sűrűsége minták és struktúrák modellezésére és felismerésére használatos a jelfeldolgozások során. A kölcsönösségi együtthatók központi szerepet játszanak a jelfeldolgozásban, a telekommunikációs rendszerekben, beleértve a prediktív kódolókat, kiegyenlítőket, digitális dekódereket, késleltetés kiszámítókat, osztályozókat és jelvisszaállító rendszereket. Egy  $X(m)$  véletlenszerű folyamat autokorrelációs függvénye, melyet  $r_{xx}(m_1, m_2)$ -vel jelölünk, a következőképpen definiálható:

$$r_{xx}(m_1, m_2) = E[x(m_1)x(m_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(m_1)x(m_2) f_{X(m_1), X(m_2)}(x(m_1), x(m_2)) dx(m_1) dx(m_2)$$

Az  $r_{xx}(m_1, m_2)$ -vel jelölt autokorrelációs függvény az  $X$  folyamat végeredményének  $m_1$  és  $m_2$  időpillanatbeli hasonlóságának mértékét vagy a kölcsönösségi relációját mutatja meg. Ha a véletlenszerű folyamat végeredménye az  $m_1$  és az  $m_2$  időpillanatbeli értéke nincs hordoz relációt, akkor  $X(m_1)$  és  $X(m_2)$  függetlennek vagy korrelálatlannak mondható, tehát  $r_{xx}(m_1, m_2) = 0$ . A nem szigorú értelemben vett stacionárius folyamatok esetében az autokorrelációs függvény idő invariáns és a két időpillanat között eltelt idő hosszának függvénye, vagyis  $m = m_1 - m_2$ :  $r_{xx}(m_1 + \tau, m_2 + \tau) = r_{xx}(m_1, m_2) = r_{xx}(m_1 - m_2) = r_{xx}(m)$ . Egy valós értékű nem szigorú értelemben vett stacionárius folyamat autokorrelációs függvénye szimmetrikus a következő tulajdonságokkal:  $r_{xx}(-m) = r_{xx}(m)$  és  $r_{xx}(m) \leq r_{xx}(0)$ . Ez esetben az  $r_{xx}(0)$  jelenti a nulla középértékű jel energiáját.

#### 2.2.1.4.3. Autokovariancia

Az autokovariancia függvény egy  $X(m)$  véletlenszerű folyamat  $c_{xx}(m_1, m_2)$  jelöléssel ellátott szórásának vagy szóródásának mérésére szolgál a folyamat középértéke körül és a következőképpen definiálható:

$$c_{xx}(m_1, m_2) = E[(x(m_1) - \mu_x(m_1))(x(m_2) - \mu_x(m_2))] = r_{xx}(m_1, m_2) - \mu_x(m_1)\mu_x(m_2),$$

ahol a  $\mu_x(m)$  jelenti az  $X(m)$  véletlenszerű folyamat középértékét. Nulla középértékű folyamatok esetében az autokovariancia és az autokorreláció függvények azonosak. A  $c_{xx}(m_1, m_2)$  a folyamat varianciáját (szórásnégyzetét) jelenti. Stacionárius folyamatok esetében az autokovariancia függvény az előző egyenletből a következőképpen alakul:  $c_{xx}(m_1, m_2) = c_{xx}(m_1 - m_2) = r_{xx}(m_1 - m_2) - \mu_x^2$ .

#### 2.2.1.4.4. Spektrális energiasűrűség

Egy véletlenszerű folyamat energiájának spektrális sűrűségfüggvénye, más néven energiaspektruma megadja a jel egyes frekvenciákhoz tartozó energiáinak eloszlását a teljes spektrumon. A nem szigorú értelemben vett stacionárius folyamat  $X(m)$  energiaspektruma a Wiener–Khinchin tétel alapján úgy definiálható, hogy az autokorrelációs függvény Fourier transzformáltja és a következőképpen írható fel:

$$P_{XX}(f) = E[X(f)X^*(f)] = \sum_{m=-\infty}^{\infty} r_{xx}(m) e^{-j2\pi f m},$$

ahol  $r_{xx}(m)$  és  $P_{XX}(f)$  rendre az  $x(m)$  autokorrelációs függvénye és energiaspektruma, és  $f$  pedig a frekvenciaváltozó. Egy valós értékű stacionárius folyamat autokorrelációs függvénye szimmetrikus, és az energiaspektruma a következőképpen írható fel:

$$P_{XX}(f) = r_{xx}(0) + \sum_{m=1}^{\infty} 2r_{xx}(m) \cos(2\pi f m)$$

A spektrális sűrűségfüggvény egy nemnegatív valós értékészletű felvevő függvény, megadása pedig watt per herzben  $\left[\frac{\text{W}}{\text{Hz}}\right]$  történik. Az autokorrelációs so-

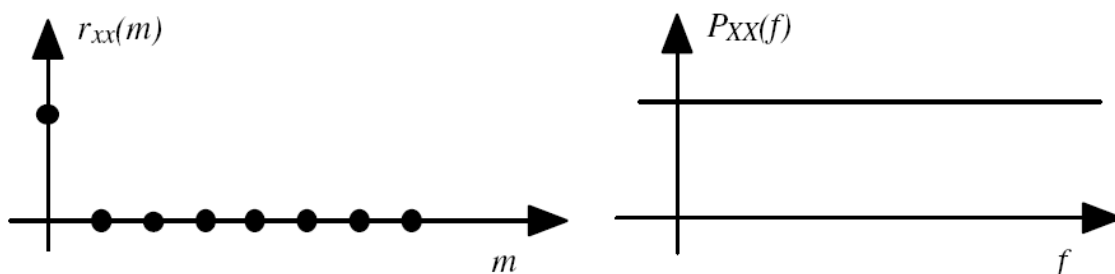
rozat a véletlenszerű folyamat energiaspektrumából az iménti egyenlet felhasználásával kapható meg az inverz Fourier transzformáció alkalmazásával a követke-

zőképpen:  $r_{xx}(m) = \int_{-\frac{1}{2}}^{\frac{1}{2}} P_{XX}(f) e^{j2\pi fm} df$ . Az autokorreláció és az energiaspektrum a

folyamat másodrendű statisztikai, rendre időben és frekvenciatartományban.

#### Példa – fehér zaj

A következő ábra a fehér zaj energiaspektrumát és autokorrelációját mutatja.



13. ábra A fehér zaj autokorrelációja és energispektruma

Egy olyan zajt, amely korrelálatlan, egymástól független mintákból áll, fehér zajnak nevezzünk. Egy stacionárius fehér zaj  $n(m)$  autokorrelációja definiálható:

$r_{nn}(k) = E[n(m)n(m+k)] = \begin{cases} \text{Zajenergia} & k = 0 \\ 0 & k \neq 0 \end{cases}$ . Ez az egyenlet egy korrelálatlan fehér

zaj definíciója szerinti matematikai állítás. A frekvenciatartományra ekvivalens leírás adható meg, az  $r_{nn}(k)$  Fourier transzformációjával a következőképpen:

$P_{NN}(f) = \sum_{k=-\infty}^{\infty} r_{nn}(k) e^{-j2\pi fk} = r_{nn}(0) = \text{zajenergia}$ . A stacionárius fehér zaj energiaspekt-

ruma azonosan terjed szét a sáv teljes szélességében és időben. A fehér zaj az egyik legnehezebben eltávolítható zajforma, mert nem határozható meg a szerkezete sem időben, sem a frekvenciasávok szerint.

#### Példa – impulzív zaj

Az impulzív zaj egy véletlenszerű kétállapotú („be/ki”) impulzussorozatból áll, véletlenszerű amplitúdóval és időben véletlenszerű előfordulással. Egy véletlenszerű impulzív zaj sorozat  $n_i(m)$  egy kétállapotú véletlenszerű amplitúdó modulált sorozatként modellezhető, a következő szerint:  $n_i(m) = n(m)b(m)$ , ahol  $b(m)$  egy

kétállapotú sorozat, mely egy impulzus jelenlétét vagy hiányát jelzi, és  $n(m)$  pedig egy véletlenszerű zaj. Tegyük fel, hogy az impulzív zaj egy korrelálatlan folyamat, így az impulzív zaj, mint egy kétállapotú folyamat autokorrelációja a következőképpen definiálható:  $r_m(k, m) = E[n_i(m)n_i(m+k)] = \sigma_n^2 \delta(k)b(m)$ , ahol  $\sigma_n^2$  a zaj varianciája. Ebben az egyenletben az autokorrelációs függvény egy kétállapotú függvényként lett kifejezve, mely az impulzív zaj  $m$  időpillanatbeli be/ki állapotától függ. Az impulzív zaj sorozat energiaspektruma az imént leírt autokorrelációs függvény Fourier transzformáltjaként kapható meg a következő összefüggéssel:  $P_{NN}(f, m) = \sigma_n^2 b(m)$ .

#### 2.2.1.4.5. Két véletlenszerű folyamat együttes statisztikája

A jelfeldolgozás során gyakran találkozunk azzal a problémával, például egy több érzékelőből álló rendszer esetében, hogy nem csak egy véletlenszerű folyamat van jelen. Az együttes statisztika és együttes eloszlás leírja a két vagy több folyamat kapcsolatát egymással. Két diszkrét időbeosztással mintavételezett véletlenszerű folyamat  $x(m)$  és  $y(n)$  együttes eloszlásfüggvénye a következő jelöléssel adható meg:  $f_{X(m_1)\dots X(m_M), Y(n_1)\dots Y(n_N)}(x_1, \dots, x_M, y_1, \dots, y_N)$ . Ha a két véletlenszerű folyamat  $X(m)$  és  $Y(n)$  korrelálatlanok, az együttes eloszlásfüggvényük a két eloszlásfüggvény szorzataként adható meg a következő felírással:

$$f_{X(m_1)\dots X(m_M), Y(n_1)\dots Y(n_N)}(x_1, \dots, x_M, y_1, \dots, y_N) = f_{X(m_1)\dots X(m_M)}(x_1, \dots, x_M) f_{Y(n_1)\dots Y(n_N)}(y_1, \dots, y_N)$$

#### 2.2.1.4.6. Kereszt-korreláció és kereszt-kovariancia

Két véletlenszerű folyamat kereszt-korrelációja az alábbiak szerint adható

$$\text{meg: } r_{xy}(m_1, m_2) = E[(x(m_1)y(m_2))] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(m_1)y(m_2)f_{X(m_1)Y(m_2)}(x(m_1), y(m_2))dx(m_1)dy(m_2)$$

Egy  $r_{xy}(m_1, m_2)$  nem szigorú értelemben vett stacionárius folyamat esetében a korrelációs függvény csak az  $m_1$  és  $m_2$  időpillanatok különbségétől függ, vagyis  $m = m_1 - m_2$ , tehát:  $r_{xy}(m_1 + \tau, m_2 + \tau) = r_{xy}(m_1, m_2) = r_{xy}(m_1 - m_2) = r_{xy}(m)$ .

A kereszt-kovariancia függvény a következőképpen definiálható:  $c_{xy}(m_1, m_2) = E[(x(m_1) - \mu_x(m_1))(y(m_2) - \mu_y(m_2))] = r_{xy}(m_1, m_2) - \mu_x(m_1)\mu_y(m_2)$ . Nulla középértékkel rendelkező folyamatok esetében a kereszt-korreláció és a kereszt-kovariancia függvények azonosak. Nem szigorú értelemben vett stacionárius folyamatok esetében a kereszt-kovariancia függvény az előző egyenlet alapján:  $c_{xy}(m_1, m_2) = c_{xy}(m_1 - m_2) = r_{xy}(m_1 - m_2) - \mu_x\mu_y$ .

## 2.2.1.4.7. Kereszt-energiáspektrum eloszlás és koherencia

Két véletlenszerű folyamat  $X(m)$  és  $Y(m)$  kereszt-energiáspektrum eloszlása a két folyamat kereszt-korrelációs függvényének Fourier transzformációjával definiálható:

$P_{XY}(f) = E[X(f)Y^*(f)] = \sum_{m=-\infty}^{\infty} r_{xy}(m)e^{-j2\pi fm}$ . Ahogyan a kereszt-korreláció a

kereszt-energiáspektrum eloszlása is a két jel hasonlóságának mértékét mutatja, vagy a két folyamat spektrális koherenciáját, mely a kereszt-energiáspektrum eloszlásának normalizált formulája és a következőképpen definiált:

$C_{XY}(f) = \frac{P_{XY}(f)}{\sqrt{P_{XX}(f)P_{YY}(f)}}$ . A koherencia függvény olyan alkalmazásokban haszná-

latos, mint az időkésettetés kiszámítása és a jel-zaj arány meghatározása (mérése).

## 2.2.2. A rejtett Markov modell

A rejtett Markov modelleket a véletlenszerű nem stacionárius jelek statisztikai modellezésére használhatjuk hatékonyan. Az ilyen jelek csoportjába tartozik az emberi beszéd vagy az időben változó zaj. A HMM egy Markov láncot ad meg, mely az időbeli változást vagy a terek változását állapotfüggő stacionárius alfolyamatokra bontja. Egy HMM lényegében egy véges állapothalmazzal rendelkező Bayes folyamat, Markov előzménnyel, az egyes állapotok között átmenetek modellezésére, és egy halmaz valószínűségi eloszlásfüggvény a jel véletlenszerű variációinak modellezésére minden állapoton belül.

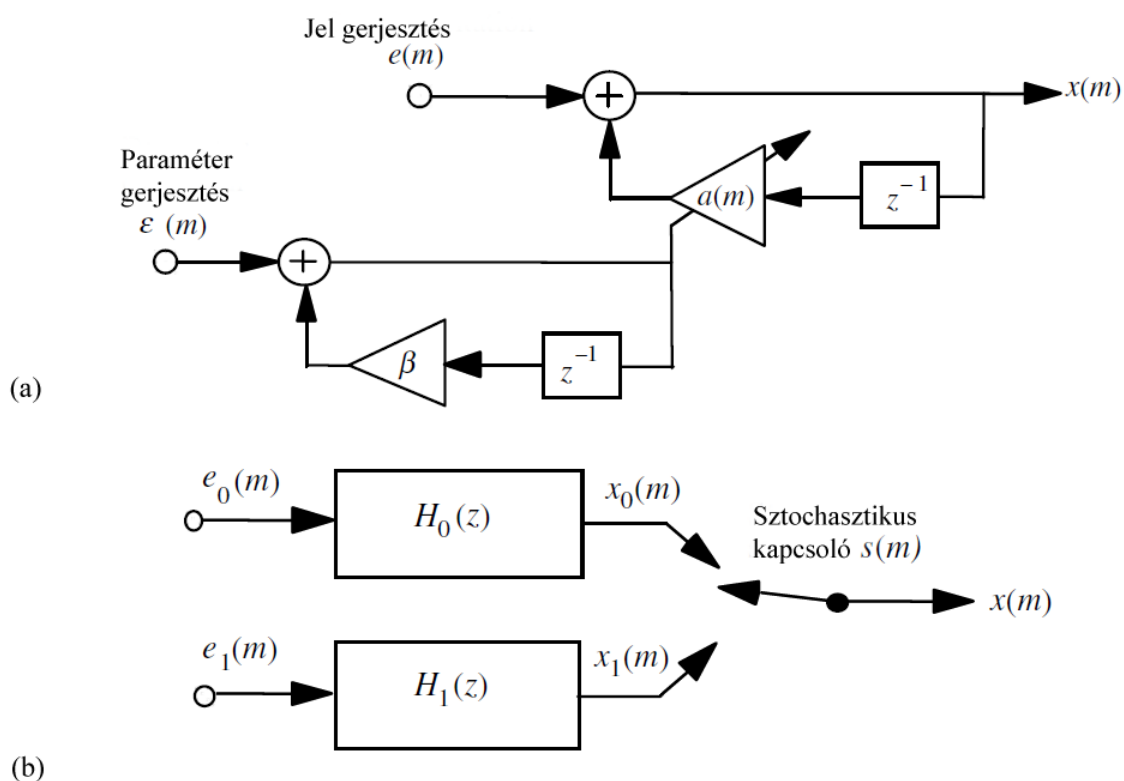
2.2.2.1. Nem stacionárius folyamatok statisztikai modellje

Egy nem stacionárius folyamat definiálható oly módon, hogy egy olyan folyamat, melynek statisztikai paraméterei az idővel változnak. A legtöbb „természetesen előállított” jel, mint például a hangjelek, képek, biomedikai és szeizmikus jelek nem stacionárius jelek, mivel az azokat előállító rendszerek paraméterei és a környezet, melyben terjednek változnak az idő múlásával.

A nem stacionárius jelek egy kétrétegű sztochasztikus folyamatként modellezhetők, egy rejtett folyamattal az észlelhető jel statisztikai paramétereinek időbeli változásainak kontrollálására. Általában a nem stacionárius folyamatok két alapvető nagy csoportba sorolhatók:

- a.) folytonosan változó állapotú folyamatok
- b.) véges sok állapotú folyamatok

A folytonosan változó állapotú folyamatok definíciós szerint az időben folytonosan változó jelek statisztikájának alapjául szolgálnak. Ilyenek például az audiojelek, a beszéd és a zene, melyeknek az energiája és spektrális összetétele az idővel változik. Véges sok állapottal rendelkező folyamat az olyan statisztikai jellemzőkkel rendelkező folyamat, mely képes váltani a véges sok stacionárius vagy nem stacionárius állapot között. Például az impulzív zaj egy ilyen kétállapotú folyamat. Folytonosan változó folyamatok közelíthetők hozzájuk illő véges sok állapotú folyamattal. A következő ábra egy nem stacionárius autoregresszív (AR) folyamatot mutat be. Ez a folyamat a jel paramétereire egy rejtett stacionárius AR modell és magára a jelre egy időben változó AR modell kombinációját alkalmazza. A



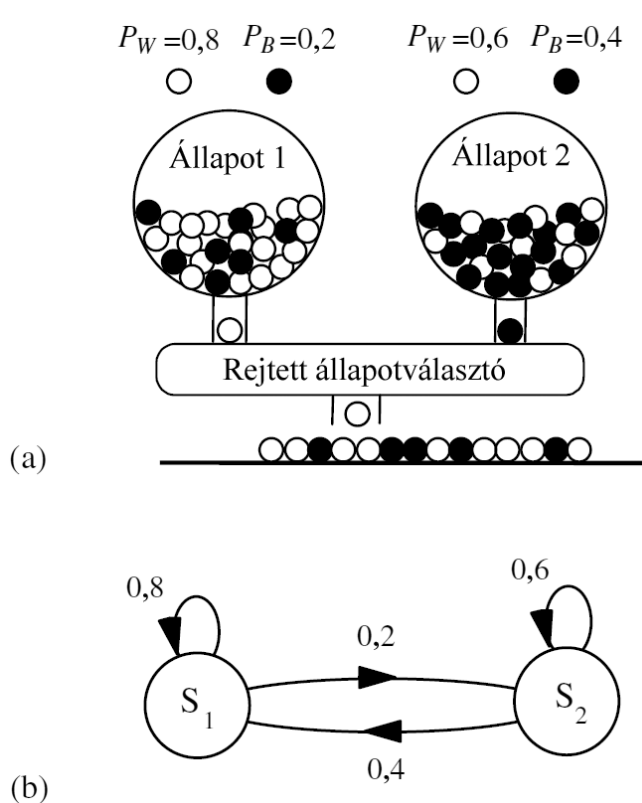
14. ábra Folytonosan változó (a) és kétállapotú (b) AR folyamat

rejtett modell vezérli a nem stacionárius AR modell időben változó paramétereit. Ehhez a modellhez a megfigyelt jel egyenlete és a paraméter állapotegyenlete a következő módon írható fel: a megfigyelt jel  $x(m) = a(m)x(m-1) + e(m)$  és a rejtett állapot  $a(m) = \beta a(m-1) + \varepsilon(m)$ , ahol  $a(m)$  a megfigyelt AR folyamat időben változó együtthatója és  $\beta$  pedig a rejtett állapotirányító folyamat együtthatója. A 14. (b) ábra egy egyszerű példát mutat a véges sok állapotú nem stacionárius folyamat kétállapotú autoregresszív megvalósítására, amelyben egy közvetlen kapcsoló választja ki minden időpillanatban a két AR modell közül az egyiket, melyet a

kimenetre csatol. Ehhez a modellhez a kimeneti jel  $x(m)$  a következő egyenlettel fejthető ki:  $x(m) = \bar{s}(m)x_0(m) + s(m)x_1(m)$ , ahol a kétállású kapcsoló  $s(m)$  kiválasztja a folyamat állapotát  $m$  időpillanatban, és  $\bar{s}(m)$  jelöli  $s(m)$  ellentettjét.

### 2.2.2.2. Rejtett Markov modellek

A rejtett Markov modell (HMM) egy kétrétegű véges sok állapotú folyamat, egy rejtett Markov folyamattal, mely irányítja az állapotok kiválasztását a figyelembe veendő folyamatban. Mint egy egyszerű illusztrációt figyeljük meg a következő ábrán látható kétállapotú Markov folyamatot, mely két tárolót mutat különböző keverékű fehér és fekete színű golyókkal. A fekete és fehér golyók előfordulásának valószínűségét jelöljük rendre  $P_B$  és  $P_W$  jelölésekkel, ahogyan az ábra mutatja. Tegyük fel, hogy folytonos idő intervallumban egy rejtett kiválasztó folyamat kivá-



15. ábra Kétrétegű véletlenszerű folyamat (a), és rejtett Markov modellje (b)

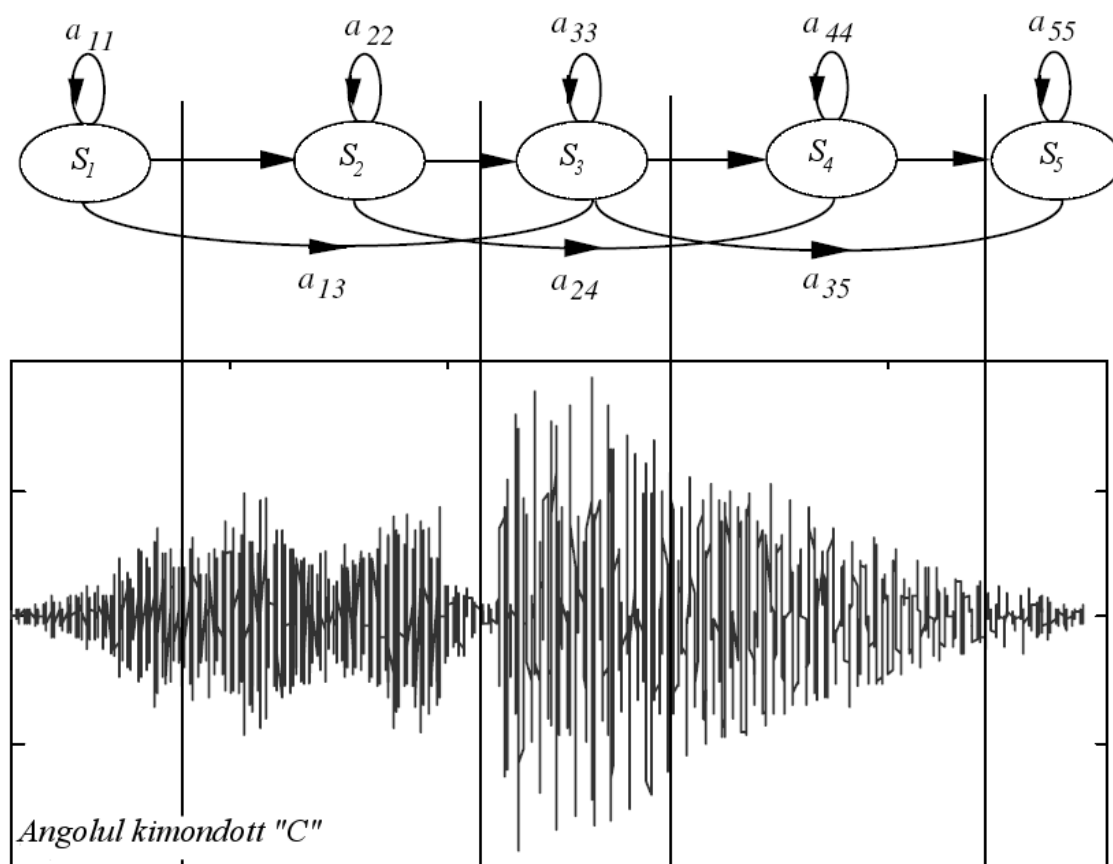
lasztja az egyik tartályt, hogy kiejtsen egy golyót. A golyót kiejtése után pótoljuk, tehát a fehér és fekete golyók számára nincs hatása a műveletnek. Minden tárolót úgy tekinthetünk, hogy az egy alapvető állapota a kimeneti folyamatnak. Most tegyük fel, hogy a rejtett tároló-kiválasztót a következő szabály irányítja: bármely időpillanatban, ha a kiválasztott tárolóból fehér golyót kapunk, akkor nem választunk másik tárolót, ha feketét akkor a másik tárolót választjuk. Ez egy példa a Markov folyamatra, mert a folyamat következő állapota függ a jelenlegi állapottól, ahogyan azt az ábra kétállapotú modellje is mutatja. Ebben a modellben a megfigyelhető végeredmény nem jelzi egyértelműen az alapul vett állapotot, mivel mindkét állapot (tároló) alkalmas minkét színű (fekete és fehér) golyó kibocsátására.

Általában a Markov modellnek  $N$  állapota van, és mindegyik állapot más szegmensét modellezi a vizsgált jelnek. A rejtett Markov modell alkalmazható egy

időben változó véletlenszerű folyamat modellezésére, mint egy  $N$  elemű stacionárius vagy kvázi-stacionárius elemi alfolyamat valószínűségi Markov láncá.

### 2.2.2.2.1. A rejtett Markov modellek fizikai értelmezése

A modell fizikai értelmezéséhez nézzük meg a következő, 16. ábrát. Az ábra balról jobbra mutatja egy angolul kiejtett „C”<sup>1</sup> hang modelljét és vele együtt a hang hullámformájának térképét is. Általában kétféle eltérés van a beszéd és az egyéb sztochasztikus jelek között: a spektrális összeállítás változása, és az időbeosztás vagy artikuláció változása. Egy rejtett Markov modellben ezek a változások állapot megjegyzéssel és az egyes állapotok közötti átmenetek valószínűségével modellezhetők. Használható módszer a rejtett Markov modell interpretálására és alkalmazására, hogy a modell minden állapotát úgy tekintjük, hogy a sztochasztikus folyamat egy szegmensét modellezi. Az ábrán az  $S_1$  állapot modellezi az első szeg-



16. ábra Ötállapotú balról–jobbra HMM beszédmodell

<sup>1</sup> Az angol abc 3. betűjének, a c-nek kiejtése a magyartól jelentősen eltér, fonetikusan a következőképpen írható le: [si:]



menet, a  $S_2$  a másodikat és így tovább. Minden állapotnak kell lennie egy olyan mechanizmusának, mely összeilleszti a szegmensek véletlenszerű variációinak különböző realizációit, melyet modelleznek. Az állapotátmenetek valószínűségei egy mechanizmust biztosítanak a különböző állapotok kapcsolatára, és a variációk modellezésére az időskála jelek alatti időtartamára minden állapotban. Például ha a beszéd egy szegmensének vége kinyúlt, köszönhetően a lassú artikulációnak, hozzáigazítható lesz több önmagába visszatérő átmenettel az állapotához, mely a szegmenst modellezi. Ellenben ha a szegmens elhallgatott a gyors beszédnek köszönhetően, akkor a következő állapotra ugrással illeszt ebben az esetben. Az állapot tapasztalati eloszlásfüggvényei modellezi a jel állapotokhoz rendelt szegmenseinek spektrális összeállításának valószínűségi eloszlását.

#### 2.2.2.2.2. A rejtett Markov modell paraméterei

A rejtett Markov modellnek a következő paraméterei vannak:

- Az állapotok száma  $N$ . Ezt általában beállítjuk az egymástól különböző vagy elemi sztochasztikus események számára egy jelfolyamatban. Például egy impulzív zajhoz hasonló kétállapotú folyamat modellezése esetén  $N$  értéke 2 lesz, míg szavankénti beszéd esetében  $N$  értékét 5 és 10 közé szokás választani.
- Állapot-átmeneti valószínűségi mátrix  $\mathbf{A} = \{a_{ij}, i, j = 1, \dots, N\}$ . Ez biztosít kapcsolatot vagy Markov hálózatot hoz létre az egyes állapotok között, és modellezi a variációkat az állapotokhoz rendelt jelek időtartama alatt. A balról–jobbra rejtett Markov modell esetében  $a_{ij} = 0$ , minden  $i > j$  esetén, ezért ez az  $\mathbf{A}$  mátrix egy felső háromszög mátrix.
- Állapot tapasztalati vektorok  $\{\mu_{i1}, \mu_{i2}, \dots, \mu_{iM}, i = 1, \dots, N\}$ . Minden állapotra egy halmaz  $M$  prototípus vektor, a jel terének centroidjainak modellezésére minden állapotban.
- Állapot tapasztalati vektor valószínűségi modellje. Lehet diszkrét az  $M$  prototípusvektorokból, azok hozzárendelt sűrűségfüggvényével  $\mathbf{P} = \{P_{ij}(\cdot); i=1, \dots, N, j=1, \dots, M\}$ , vagy folytonos (általában normális)  $\mathbf{F} = \{f_{ij}(\cdot); i=1, \dots, N, j=1, \dots, M\}$  eloszlásfüggvény modell.
- Kezdőállapot valószínűségi vektor  $\boldsymbol{\pi} = [\pi_1, \pi_2, \dots, \pi_N]$

2.2.2.3. A rejtett Markov modell betanítása

Az első lépés, hogy adatbázist alakítsunk ki a modell számára, nagy számú minta vételezésével a vizsgálat tárgyát képező jelből. Tegyük fel, hogy az adatbázis  $L$  vektor értékekkel bíró szekvenciából áll  $[\mathbf{X}] = [\mathbf{X}_k; k=0, \dots, L-1]$ , és minden szekvenciában  $\mathbf{X}_k = [x(t); t=0, \dots, T_k-1]$  változó számú  $T_k$  vektor van. A paraméterek megfelelő beállításához a modell statisztikájának adathalmazát kell előállítani. Valószínűségi értelemben a modell alkalmassága a modell utólagos valószínűségével  $P_{M|X}(M|\mathbf{X})$  adható meg,  $M$  modell  $\mathbf{X}$  adathalmazára. Ez a betanító folyamat az  $M$  modell utólagos valószínűségének és az  $[\mathbf{X}]$  adathalmaznak a maximalizálását célozza, a Bayes szabály alkalmazásával:  $P_{M|X}(M|\mathbf{X}) = \frac{1}{f_X(\mathbf{X})} f_{X|M}(\mathbf{X}|M) P_M(M)$ ,

ahol a nevezőben álló  $f_X(\mathbf{X})$  egy normalizáló hatással bír, és  $P_M(M)$  az  $M$  modell korábbi valószínűsége. Az iménti egyenlet maximalizálása egyenértékű a valószínűségi függvény maximalizálásával, mely a következő egyenlet segítségével fejezhető ki  $\mathbf{X}$  vektorszekvenciára és  $M$  modellre:

$$f_{X|M}(\mathbf{X}|M) = \sum_s f_{X|S,M}(\mathbf{X}|s, M) P_{s|M}(s|M),$$

ahol  $f_{X|S,M}(\mathbf{X}|s, M)$  jelöli az eloszlásfüggvényt az  $M$  modell  $\mathbf{X}$  szekvenciájának  $\mathbf{s} = [s(0), s(1), \dots, s(T-1)]$  állapotában, mely a következőképpen fejezhető ki:

$$f_{X|S,M}(\mathbf{X}|s, M) = f_{X|S}(\mathbf{x}(0)|s(0)) f_{X|S}(\mathbf{x}(1)|s(1)) \cdots f_{X|S}(\mathbf{x}(T-1)|s(T-1)),$$

ahol az  $N$  állapot egyike lehet  $s(t)$   $t$  időpillanatban,  $f_{X|S}(\mathbf{x}(t)|s(t))$  pedig

$f_{X|S,M}(\mathbf{x}(t)|s(t), M)$  rövidített leírása az  $M$  modell  $\mathbf{x}(t)$  eloszlásfüggvényének  $s(t)$

állapotban. A  $\mathbf{s}$  állapotszekvencia Markov valószínűsége a következőképpen adható

$$\text{meg: } P_{s|M}(s|M) = \pi_{s(0)} a_{s(0)s(1)} a_{s(1)s(2)} \cdots a_{s(T-2)s(T-1)}.$$

Az előző három egyenletből az utolsó kettőt behelyettesítve az elsőbe kapjuk a következőt:

$$\begin{aligned} f_{X|M}(\mathbf{X}|M) &= \sum_s f_{X|S,M}(\mathbf{X}|s, M) P_{s|M}(s|M) = \\ &= \sum_s \pi_{s(0)} f_{X|S}(\mathbf{x}(0)|s(0)) a_{s(0)s(1)} f_{X|S}(\mathbf{x}(1)|s(1)) \cdots a_{s(T-2)s(T-1)} f_{X|S}(\mathbf{x}(T-1)|s(T-1)) \end{aligned}$$

ahol az összegzés végighalad a összes  $\mathbf{s}$  állapotszekvenciát. A betanítási folyamatban az átmeneti valószínűségeket és a tapasztalati eloszlásfüggvények paramétereit az úgy kell kiszámítani, hogy iménti egyenletből a modell valószínűsége maximális legyen. A közvetlen maximalizálás a modell paramétereivel összefüggésben nem triviális megoldás. Következésképpen az egyenlet számításigénye egy megfigyelési

---

szekvencia  $T$  darab vektorára  $O(N^T)$  lesz. Ez még kis számok esetében is nagyon nagy számításigényt jelent. Az állapot-idő diagram ismétlődő struktúrái azt mutatják, hogy sok az ismétlődő számítás az egyenletben, melyet egy hatékony megvalósítással el tudunk kerülni.

### 3. Zajsűrési módszerek

#### 3.1. Zajzár

Érdekes, bár nem túl hatékony megoldás a jelszint folyamatos vizsgálata, és ha az elér egy kellően alacsony értéket, akkor a jelet teljesen kikapcsolva, csendet kapunk eredményül. Ezzel nem csak a zajt állítjuk meg, hanem az alacsony jelszintű hasznos jelet is, és a jel zajterhelése továbbra is megmarad, ha már egy adott szint felett van az értéke. Ez a módszer megoldást jelenthet olyan esetekben, mikor egy berendezés öngerjesztő folyamatának szeretnénk gátat szabni. Ilyen előfordulhat például mikrofonok esetén, melyek egy légtérben vannak az erősített jelüket megszólaltató hangsugárzókkal, és elhelyezéssel már nem tudunk a problémán segíteni (nagy hangerőre van szükség, és bárhová helyezzük a mikrofont, egy zárt térben, beindul az oszcilláció). Másik szemléletes példa az elektromos gitárok torzító berendezése, melynek alapvető működéséhez tartozik az átlagosnál nagyobb erősítés. Ilyenkor az érkező jel kikapcsolása a húrok öngerjesztő rezgésének megakadályozását szolgálhatja. Egy ilyen berendezés megfelelő beállítása nehéz, főleg, ha több vagy többféle hangforrás is van, ráadásul a valódi problémára nem jelent megoldást.

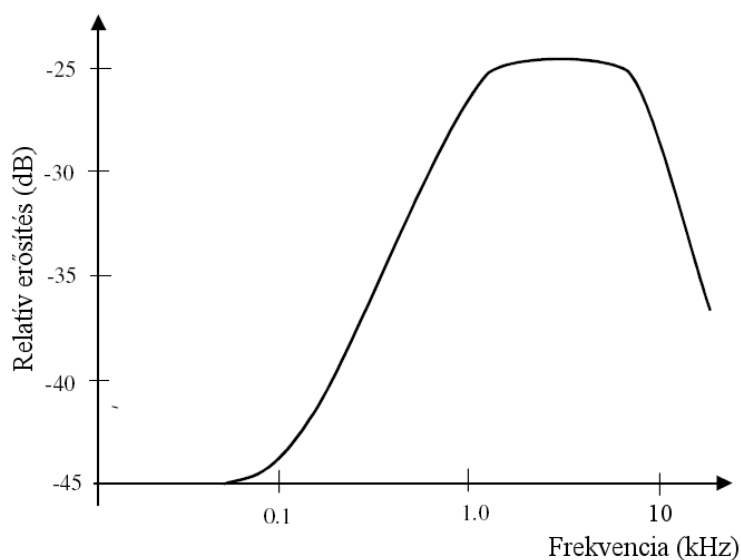
#### 3.2. Analóg valós idejű zajsűrők

##### 3.2.1. Dolby zajsűrő

A Dolby zajsűrő rendszerek úgy működnek, hogy emelik a magasabb frekvenciákon a hasznos jel amplitúdóját, ezzel javítva a jel–zaj arányt ezen frekvenciákon. A jel nagy része főként az alacsonyabb frekvenciákra – 2 kHz alá – koncentrálódik. A magasabb frekvenciájú jelek, melyek a minőséget és az élményt hordozzák viszonylag alacsonyabb energiájúak, és tovább degradálódnak a zajjal. Például egy mágneses szalagra rögzített jelsorozat esetében a szalag zaja és a korrekciós előerősítő fokozat zaja jelentős magasabb frekvenciájú „sistergést” állít elő. Lejátszásakor ezért ezeken a frekvenciákon romlik a jel–zaj arány – tehát nem lesz egyenletes a teljes átvitt sávban, vagyis az alacsonyabb frekvenciákon jobb, a magasabban rosszabb értéket kapunk. Tehát a magasabb frekvenciákon kevesebb hallható hasznos jelet kapunk. A Dolby zajsűrő tehát azon az alapelven működik, hogy a viszonylag kisebb energiájú magasabb frekvenciájú jeleket kiemeli, erősíti a hanganyag felvétele során, a többi jelet pedig változatlanul hagyja. Felvétel közben a jel

egy kiemelés és egy dinamika kompresszió keverékén megy keresztül, melyet lejátszáskor vissza kell alakítani az eredeti állapotába. A kódoló és a dekódoló rendszernek jól illesztettnek kell lennie, a torzítások elkerülése érdekében.

A Dolby laboratóriumban többféle ilyen szűrőt fejlesztettek ki, melyek a Dolby A, Dolby B, és Dolby C. Ezek alapjában véve a sávok számában és az alkalmazott kiemelő stratégiájában különböznek egymástól. A Dolby A a professzionális felhasználás számára lett kifejlesztve, és négy nagy részre osztja a frekvenciaspektrumot – az első sáv 0–80 Hz (alul áteresztő szűrő), a második sáv 80–3000 Hz (hullámsávszűrő), a harmadik sáv 3 kHz körüli felül áteresztő szűrő, a negyedik sáv pedig egy 9 kHz körüli felül áteresztő szűrő segítségével jött létre. A kódoló minden hullámsáv erősítését alkalmazkodóan igazította, az alacsony energiájú jelek kiemelése érdekében. A Dolby A rendszer gondoskodik a 10–15 dB-es erősítésről minden sávban, ha a jel 45 dB-nél nagyobb értékkel esik a maximális rögzítési jelszint alá. A kereskedelmi forgalomba kerülő készülékekben a Dolby B és a Dolby C zajsűrőket alkalmazták. Ezek a rendszerek az imént ismertetett Dolby A rendszerrel szemben már csak két hullámsávon dolgoznak, a négy helyett. A



17. ábra A Dolby C rendszer elő-kiemelő sávátvitele

Dolby B rendszer 10 dB-es emelést biztosít alacsony jelszint esetén (nagyobb, mint 45 dB eltérés a maximális rögzítési jelszinthez képest), a Dolby C rendszer pedig 20 dB-es emeléssel dolgozik, ahogy az a következő ábrán látható.

Ez a módszer csak analóg jelek esetében egy jól meghatározott zajforma csökkentésére alkalmas,

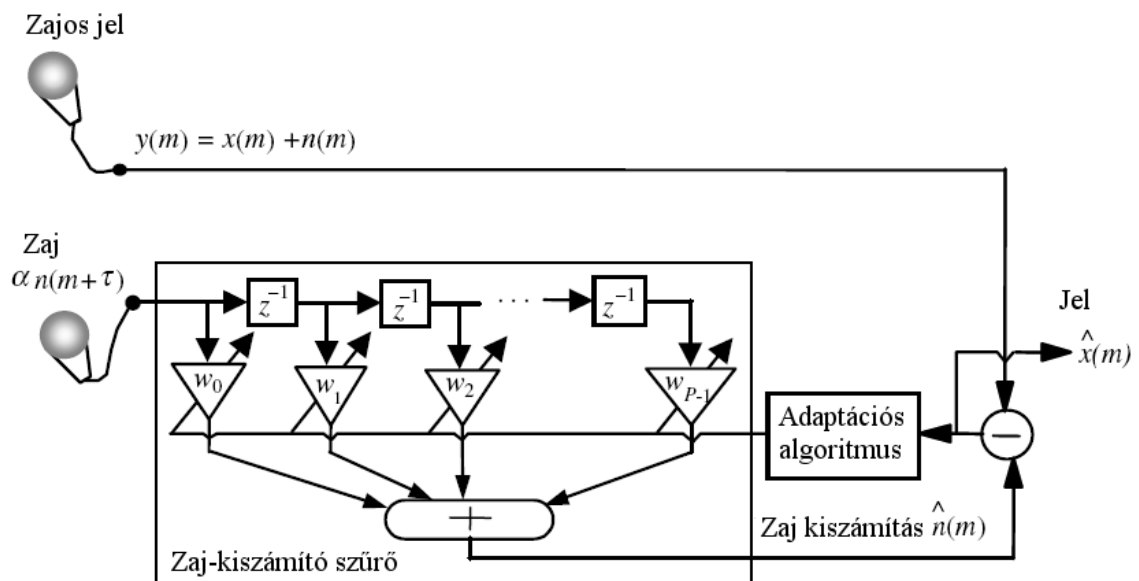
ezért ez nagyon behatárolja az alkalmazási területet. Ezért alkalmazzák ezt a módszert szalagos rögzítő eszközök – úgy mint kompakt kazettás magnetofonok – esetében. Az ilyen készülékek esetében jellemző, hogy a szalagon található mágneses részecskék rendezettsége és az előmágnesezés milyensége befolyásolja a hangminőséget, továbbá a magasabb frekvenciájú hangjelek – az alkalmazott mágne-

szehető anyag függvényében – kisebb jelszinten érkeznek vissza a szalagról, mely egy korrekciós előerősítő fokozat alkalmazását teszi szükségessé, ezáltal az elektronikus rendszernek is lesz egy zaja, melyek együttesének „szűrésére” alkalmazható a Dolby A, B és C rendszer.

### 3.2.2. Zajérvénytelenítés

Egy zajos környezetből (mint például egy mozgó autó) származó beszéd esetén a hasznos jelet a környezet zajával együtt észleljük. A mérőrendszerek esetében az információhordozó jel gyakran a berendezést körülvevő környezet zajával szennyezett. A zajjal együtt észlelt  $y(m)$  jel a következőképpen modellezhető:  $y(m) = x(m) + n(m)$ , ahol  $x(m)$  jelenti a hasznos jelet,  $n(m)$  jelenti a zajt,  $m$  pedig a diszkrét idő mérőszámot. Néhány esetben, például mikor mobiltelefont használunk egy mozgó autóban, vagy rádiókommunikációs eszközt egy repülőgép pilótafülkéjében, talán lehetséges, hogy kimérhető és kiszámítható a környezeti zaj pillanatnyi amplitúdója, egy irányított mikrofon használatával. Ezután a az  $x(m)$  hasznos jel visszaállítható lehet a zajos jel és a zaj különbségeként.

Az adaptív zaj-érvénytelenítő rendszer – két bemenettel rendelkezik, melyekből az egyik a környezeti zajt érzékeli, a másik pedig a zajjal terhelt hasznos jelet – alkalmazható beszédhang kiemelésre. Ebben a rendszerben az egyik irányított mikrofon a zajos  $x(m) + n(m)$  jelet érzékeli, míg a másik irányított mikrofon az előzőtől egy adott távolságra elhelyezve a zajt  $\alpha \cdot n(m + \tau)$  méri. Az a csillapítási szor-



18. ábra Adaptív zajérvénytelenítő elvi felépítése

zótényező és a  $\tau$  időkéleltetési érték meglehetősen túlegyszerűsített modellt ad a mikrofon elhelyezésétől eltérő helyről származó és eltérő zajok hatásainak modellezésére. A második mikrofonból érkező zajt egy adaptív digitális szűrővel dolgozzuk fel, annak érdekében, hogy amplitúdója megegyezzen az első mikrofonban érzékelt hang zajszenyezésének amplitúdójával, ezután pedig ezt a jelet kivonjuk zajos jelből, a zaj érvénytelenítése céljából. Ez a berendezés a zaj alacsonyabb frekvenciáin működik hatékonyan, de szenved a jelek nem állandó jellegétől, és a túlegyszerűsített feltevéstől mely szerint lineáris szűrővel modellezhető a zaj szóródása és terjedése a térben. A berendezés felépítését a 18. ábra szemlélteti.

Sok esetben, például egy telekommunikációs rendszerben alkalmazott vevőkészülék esetében nem ismerjük a zajszenyezés pillanatnyi értékét, csak a zajos jel adott. Ezen esetekben nem érvényteleníthető a zaj, csak mértéke csökkenthető, a jel és a zaj statisztikáját felhasználva. A Wiener szűrő alkalmas ilyen jellegű zajcsökkentési eljárásra, ahol csak a zajos jel adott.

### **3.3. Digitális jelek feldolgoása**

#### *3.3.1. Átalakítások*

##### 3.3.1.1. Analóg–digitális átalakítás

Az analóg jelfeldolgozás során valamilyen folytonosan változó fizikai mennyiséget alakítunk elektronikusan feldolgozható jellé. Ilyen átalakítás történik például egy mikrofonban, ahol a levegő rezgéseit, nyomásváltozásait alakítjuk elektromos feszültséggé, melyet a későbbiekben elektronikus jelerősítő berendezésekkel tudunk tovább feldolgozni. Ezek a jelek még elég sérülékenyek, könnyen terhelődhetnek zajjal, és ezen zajok eltávolítása az analóg jelekből nagyon nehéz feladat. Ha a jelet szeretnénk megszabadítani az additív zajoktól, akkor az első lépés, amit meg kell tennünk a jel átalakítása digitális jellé. A digitális jel fontos jellemzője, hogy egy előre meghatározott értékhalmból vehet fel csak értékeket. Ezen jelek esetében a feldolgozás már matematikai modellek alapján digitális számítógépek segítségével történhet. Továbbá nagyon előnyös tulajdonságaik közé tartozik, hogy sokkal kevésbé érzékenyek a külső additív zajokra, a diszkrét értékek miatt. A legegyszerűbb ilyen digitális jelek kétértékűek, vagy más néven bináris jelek. A mai korszerű kommunikációs rendszerekben gyakran alkalmaznak olyan készülékeket, melyek az analóg jelet először csak digitálissá alakítják, majd ez a digitális jelet

alakítják tovább egy bitfolyammá, melynek értelmezése a vevő oldalon történik. Ez a bit soros adatátvitel, napjaink egyik legkedveltebb átviteli formája.

Az analóg–digitális átalakítás vagy konverzió (Analog to Digital Conversion – ADC) úgy történik, hogy egy analóg feszültségjelet alapul véve abból adott időközönként mintát veszünk. Ez a digitális jel egyik fontos jellemzője, a *mintavételezési frekvencia*. A legjobb minőségű digitális jelet akkor kapnánk, ha egy időegység alatt végtelen sokszor vennénk mintát az analóg jelből, de ez technikailag kivitelezhetetlen és értelmetlen is. A mintavételezés a gyakorlatban rövid időközönként történik:

$T_s = \frac{1}{F_s}$ , ahol  $T_s$  az időközt,  $F_s$  pedig a frekvenciát jelöli. Ez a *mintavételezési eljárás*

az  $x(t)$  analóg jelet egy  $x(m)$  impulzussorozattá alakítja.

Az átalakítás következő lépése a kvantálás, amikor a mintavételezés után kapott impulzusokat egy diszkrét értékhalmazra képezzük. A halmaz elemeit egy analóg jeltartomány meghatározott számú részekre osztásával kapjuk meg. A kvantálás során a mintavételezés eredményeképpen kapott impulzussorozat eleméről eldöntjük, hogy melyik két érték közé (melyik tartományba) esik, és előállítjuk a  $x(m)$  jelből az  $y(m)$  jelet, mely a kvantálás eredményeképpen keletkezik. Ez a digitális jel másik fő jellemzője, amit *kvantálási hossz*nak mondunk. A jel pontossága és a visszaalakításkor kapott eredmény nagymértékben függ a tartományok számától, így a kvantálási hosszról.

A következő lépés a kódolási művelet. Ebben a műveletben minden mintavételi pontban felvett jelhez a kvantálás végeredményének megfelelő bináris kódszót rendelünk hozzá. Ezzel az eljárással minden  $y(m)$  jelhez bináris kódszót rendeltünk, tehát végeredményként kaptunk egy digitálisan feldolgozható jelet.

### 3.3.1.2. Digitális–analóg átalakítás

A digitális jel analóg jellé történő alakításakor (Digital to Analog Conversion – DAC) először a vett bináris jelsorozatot kell helyreállítani, amennyiben az szükséges. A jelsorozat visszaállítását követően a kódszavakból előállítható az  $y'(m)$  függvény, melyet egy alul-áteresztő szűrővel szűrve megkapható az eredeti  $x(t)$  jelet közelítő  $x'(t)$  analóg jel.

A közelítés javítására alkalmazhatunk interpolációt, mely a két függvény hasonlóságát növeli és csökkenti a torzítást. Az analóg jel mintavételezett formájából történő visszaállítás feltétele a Nyquist elméletben megadott. A Nyquist elmélet azt mondja ki, hogy egy sávkorlátozott jel esetében annak  $F_c$  legnagyobb frekvenci-



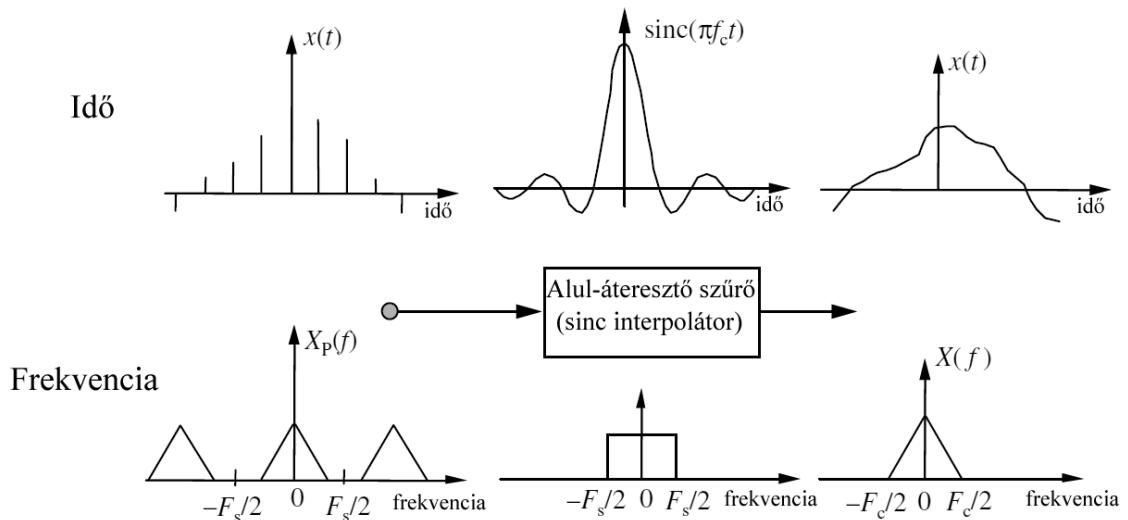
ájú összetevője is visszaállítható abban az esetben, ha a mintavételezés másodpercenkénti frekvenciája nagyobb volt a frekvencia kétszeresénél, vagyis  $F_s > 2F_c$ . Tegyük fel, hogy egy sávkorlátozott  $x(t)$  folytonos jel  $F_s$  másodpercenkénti mintavételezési frekvenciával mintavételezett. A diszkrét időközönként mintavételezett jel  $x(m)$  a következő szorzattal fejezhető ki:  $x(m) = x(t)p(t) = \sum_{m=-\infty}^{\infty} x(t)\delta(t - mT_s)$ , ahol

$p(t) = \delta(t - mT_s)$  a mintavételező függvény, és  $T_s = \frac{1}{F_s}$  a mintavételezési idő intervallum hossza. Vegyük az iménti egyenletet Fourier transzformálva, mely megmutatja, hogy a mintavételezett jel spektruma a következő egyenlettel írható fel:

$X_s(f) = X(f) * P(f) = \sum_{k=-\infty}^{\infty} X(f + kF_s)$ , ahol  $X(f)$  az  $x(t)$  jel spektruma, és  $P(f)$  a  $p(t)$

mintavételező függvénye, a  $*$  pedig a konvolúciós szorzás műveletét jelöli. Az iménti egyenlet állítása szerint a mintavételezett jel spektruma az alapsáv  $X(f)$  spektrumából és annak egyenletes  $T_s = \frac{1}{F_s}$  közönkénti ismétléséből vagy leképezéséből

áll össze. Ha a mintavételezési frekvencia felette van a Nyquist aránynak, akkor az alapsáv  $X(f)$  spektruma nincs átfedésben a  $X(f \pm kF_s)$  képeivel, és az eredeti jel



19. ábra Folytonos jel visszaállítása mintáiból. A frekvenciatartomány interpolációja az alul-áteresztő szűrővel ekvivalens

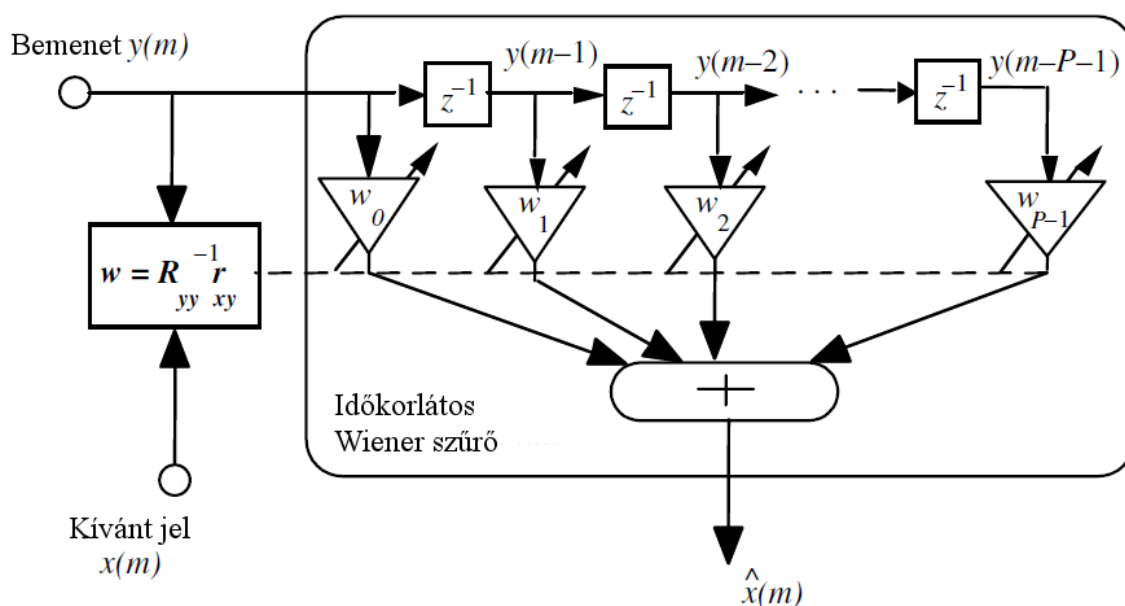
A gyakorlatban az  $F_s$  mintavételezési frekvencia nagyobb mint  $2F_s$ , általában  $2,5F_s$ , annak érdekében, hogy az átmeneti sáv szélességet hozzáigazítsuk az interpoláció alul-áteresztő szűrőjéhez.

### 3.3.2. A Wiener szűrő

A Wiener elméletet kidolgozójáról, Norbert Wiener-ről nevezték el, aki megfogalmazta az adatfüggő lineáris legkisebb négyzetek módszerére épülő szűrő elméletét. A Wiener szűrő központi szerepet játszik alkalmazásainak széles körében, mint például a visszhang érvénytelenítés, jelvisszaállítás, csatorna kompenzáció és rendszerazonosítás. A szűrő együtthatóit úgy kell kiszámítani, hogy a szűrő kimenő jele a lehető legjobban hasonlítson – vagyis különbségük négyzetének várható értéke legyen minimális – egy ismert úgynevezett megkívánt jelre. A Wiener elmélet feltételezi, hogy a jelek állandóak. Azonban ha a szűrő együtthatóit időnként újra kiszámítjuk, minden  $N$  elemű jelblokk után, a szűrő alakítja önmagát, a blokkokon belüli átlagos jellemzőkhöz, így blokk-alkalmazkodóvá válik. A blokk-alkalmazkodó (vagy szegmens-alkalmazkodó) szűrő alkalmazható olyan jelek esetében, mint például az emberi beszéd.

#### 3.3.2.1. A legkisebb négyzetek becslése

Wiener eredeti munkájában a folyamatos idejű legkisebb négyzetek módszerét adta meg, interpolációval. A módszert később kiterjesztette diszkrét időre is, mely egyszerűen megoldható, és a gyakorlatban jobban hasznosítható a digitális jelpro-



20. ábra Wiener szűrő felépítése

cesszorokban. A szűrő lehet végtelen hosszú jel feldolgozását célzó, vagy véges jel-hosszúságot figyelembevevő. Általában a végtelen szűrőnek egy halmaz nem lineáris egyenlet az eredménye, a véges hosszúságú jelre alkalmazott szűrőnek pedig egy halmaz lineáris egyenlet, mely egy zárt alakú megoldás. A véges szűrő viszonylag egyszerűbb számítást tesz lehetővé, működéséből adódóan stabilabb és praktikusabb is. Hátránya a végtelen szűrővel szemben, hogy a kívánt eredmény elérése érdekében nagyon sok együttható szükséges.

A 20. ábrán szemléltetett Wiener szűrő egy  $\mathbf{w}$  együttható vektorral megvalósított. A szűrő kap egy  $y(m)$  bemeneti jelet, és készít egy  $\hat{x}(m)$  kimeneti jelet, mely a legkisebb négyzetes eltérést mutatja az  $x(m)$  megkívánt vagy céljelhez képest. A szűrő bemenete és kimenete közötti összefüggés a következő egyenlettel írható fel:

$$\hat{x}(m) = \sum_{k=0}^{P-1} w_k y(m-k) = \mathbf{w}^T \mathbf{y}, \text{ ahol } m \text{ a diszkrét idő mérőszám, a szűrő bemeneti}$$

jele:  $\mathbf{y}^T = [y(m), y(m-1), \dots, y(m-P+1)]$ , és a paramétervektor (a Wiener szűrő együttható vektora)  $\mathbf{w}^T = [w_0, w_1, \dots, w_{P-1}]$ . A szűrési művelet kétféle egymással ekvivalens alakban fejthető ki, konvolúciós összegként, és a belső vektorok szorzataként. A hibajel  $e(m)$  a kimenőjel  $\hat{x}(m)$  és a megkívánt jel  $x(m)$  különbségeként számítható ki a következő összefüggéssel:  $e(m) = x(m) - \hat{x}(m) = x(m) - \mathbf{w}^T \mathbf{y}$ . Ebből az egyenletből kiderül, hogy egy adott  $y(m)$  bemenőjelre és  $x(m)$  megkívánt jelre a hibajel a  $\mathbf{w}$  együtthatóvektortól függ. A hibajel és az együtthatóvektor közötti összefüggés kiderítése érdekében kifejtve az előző egyenletet a következőt kapjuk:

$$\begin{pmatrix} e(0) \\ e(1) \\ e(2) \\ \vdots \\ e(N-1) \end{pmatrix} = \begin{pmatrix} x(0) \\ x(1) \\ x(2) \\ \vdots \\ x(N-1) \end{pmatrix} - \begin{pmatrix} y(0) & y(-1) & y(-2) & \dots & y(1-P) \\ y(1) & y(0) & y(-1) & \dots & y(2-P) \\ y(2) & y(1) & y(0) & \dots & y(3-P) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y(N-1) & y(N-2) & y(N-3) & \dots & y(N-P) \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_{P-1} \end{pmatrix}.$$

Ez az egyenlet  $N$  mintára,  $y(m)$  és  $x(m)$  jelekre lett felírva. Tömörített alakban a következőképpen írható fel:  $\mathbf{e} = \mathbf{x} - \mathbf{Y}\mathbf{w}$ , ahol  $\mathbf{e}$  a hibavektor,  $\mathbf{x}$  a kívánt jel vektor,  $\mathbf{Y}$  a bementi jel mátrix, és  $\mathbf{Y}\mathbf{w}$  szorzat pedig a  $\hat{x}$  kimenőjel vektor. Feltételezzük, hogy  $P$  kiinduló minta ismert, vagy hanem akkor 0 értékűnek tekintendő. Ha az előbb felírt mátrixegyenletben az  $N$  és a  $P$  értéke egyenlő, akkor egy négyzetes mátrixegyenletet kapunk, és ez az egyetlen megoldás, ahol a számítási hiba értéke nulla. Ha  $N < P$  a minták száma  $N$  elégtelen mennyiségű ahhoz, hogy elérjük az egyedülálló megoldást a szűrő együtthatóinál, ezért ebben az esetben végtelen sok megoldás létezik nulla számítási hibával és a mátrixegyenletet határozatlannak mond-

juk. A gyakorlatban a jelminták száma sokkal nagyobb, mint a szűrő hossza – vagyis  $N > P$  –, ebben az esetben a mátrixegyenletet túlhatározottnak mondjuk, és létezik egy egyedi megoldás, általában nulla számítási hibával. Mikor  $N > P$  a szűrő együtthatói úgy lesznek kiszámítva, hogy az átlagos hiba minimális legyen, mint például a hiba átlagos értékének abszolút értéke  $E[|e(m)|]$ , vagy a hiba négyzetének középértéke  $E[e^2(m)]$ , ahol  $E[.]$  a várható érték operátor. A hibafüggvény kiválasztása hatással van a megoldás optimális mivoltára, és a számítási komplexitására. A Wiener elméletben az objektív kritérium a legkisebb négyzetes hiba (LSE<sup>2</sup>) a szűrő kimeneti jele és a kívánt jel között. A legkisebb négyzetes hiba kritérium optimális a normális eloszlást követő jelek esetében. A korlátozott idő intervallumra alkalmazott legkisebb négyzetes hiba kritériumú Wiener szűrő egy lineáris, zárt formulához vezet. A Wiener szűrő együtthatói egy átlagos négyzetes hiba függvény  $E[e^2(m)]$  minimalizálásával eszközli ki, tekintettel a szűrő  $\mathbf{w}$  együttható vektorára. A számítási hiba négyzetének középértéke a következő egyenlettel adható meg:  $E[e^2(m)] = E[(x(m) - \mathbf{w}^T \mathbf{y})^2] = E[x^2(m)] - 2\mathbf{w}^T E[\mathbf{y}x(m)] + \mathbf{w}^T E[\mathbf{y}\mathbf{y}^T] \mathbf{w} = r_{xx}(0) - 2\mathbf{w}^T \mathbf{r}_{yx} + \mathbf{w}^T \mathbf{R}_{yy} \mathbf{w}$ , ahol  $\mathbf{R}_{yy} = E[\mathbf{y}(m)\mathbf{y}^T(m)]$  a bemenő jel autokorrelációs<sup>3</sup> mátrixa,  $\mathbf{r}_{xy} = E[x(m)\mathbf{y}(m)]$  pedig a bemeneti jel és a kívánt jel kereszt-kölcsönösségi vektora. Az előző kifejezést kifejtve a következő egyenletet kapjuk:

$$E[e^2(m)] = r_{xx}(0) - 2 \sum_{k=0}^{P-1} w_k r_{yx}(k) + \sum_{k=0}^{P-1} w_k \sum_{j=0}^{P-1} w_j r_{yy}(k-j), \text{ ahol } r_{yy}(k) \text{ és } r_{yx}(k) \text{ az } \mathbf{R}_{yy}$$

autokorrelációs mátrix és az  $\mathbf{r}_{xy}$  kereszt-kölcsönösségi vektor elemei külön-külön. Az előző egyenletből a korlátozott idő intervallumra alkalmazott szűrő négyzetes hibájának középérték számítására egy másodfokú függvény alkalmazható a  $\mathbf{w}$  együttható vektor értékeire, mely függvénynek egyetlen minimuma van. Például ha szűrőnek csak két együtthatója van ( $w_1, w_2$ ), akkor a négyzetes hiba középérték függvénye egy medence alakú képet mutat, egyetlen minimummal. A legkisebb négyzetes hiba jelenti az energiaminimumot.

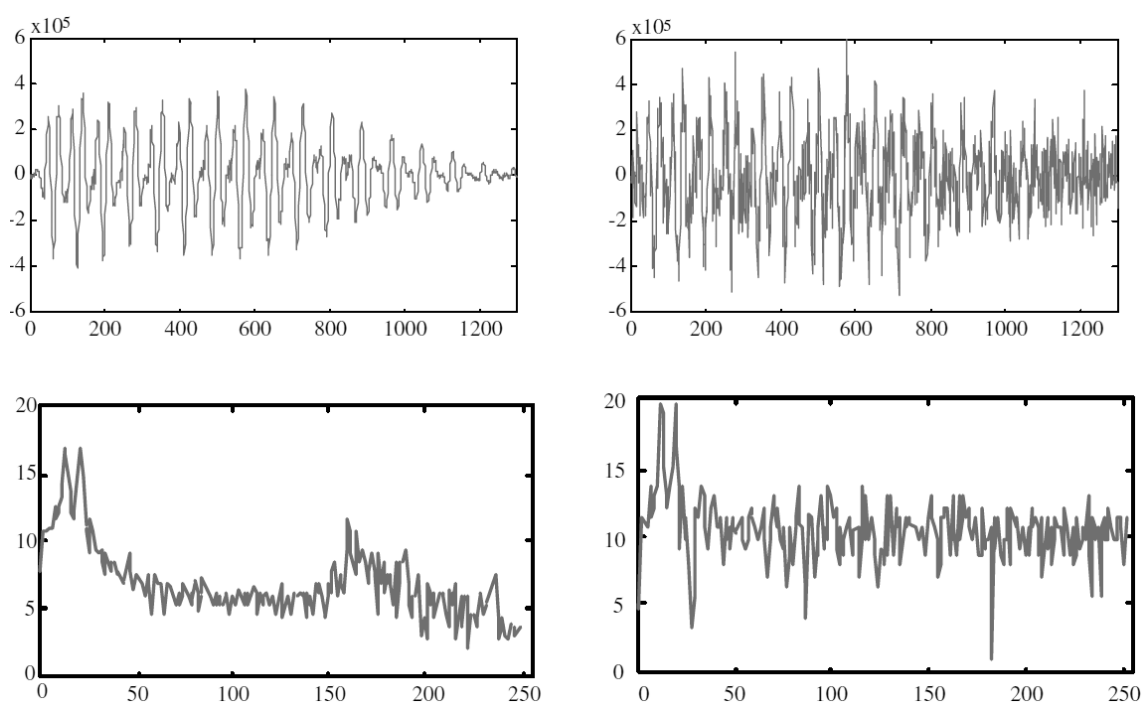
### 3.3.3. Zajcsökkentés spektrális kivonással

A spektrális kivonás módszerének lényege, hogy a zajos jelből veszünk egy olyan szakaszt, ahol a hasznos jel nincs jelen, tehát csak zaj észlelhető, és ezt a jelet egy megfelelő módszerrel kivonjuk a zajos jelből. Ez megtehető oly módon, hogy a jelek statisztikáit és spektrális paramétereit vesszük alapul. Ezzel a módszerrel adott esetben jelvisszaállítás is lehetséges.

A problémára a legegyszerűbb megoldást jelentené, ha külön rendelkezésre állna a zajos jel és a zaj. Ebben az esetben egy egyszerű művelettel eltávolítható lenne a zaj. A zajérvénytelenítő, melynek működését már korábban kifejtettem pontosan ilyen elven működik. Általában azonban csak a zajos jel áll rendelkezésre, tehát nincs lehetőség a zaj érvénytelenítésére, de előfordulhat, hogy a zaj hatását csökkenteni tudjuk a hasznos jel spektrumán.

### 3.3.3.1. A spektrális kivonás módszere

A 21. ábra olyan jeleket mutat, melyek zajjal terhelték, de nincs lehetőség a zaj érvénytelenítésére. A hasznos jel spektrumának középvértéke növelhető úgy, hogy a

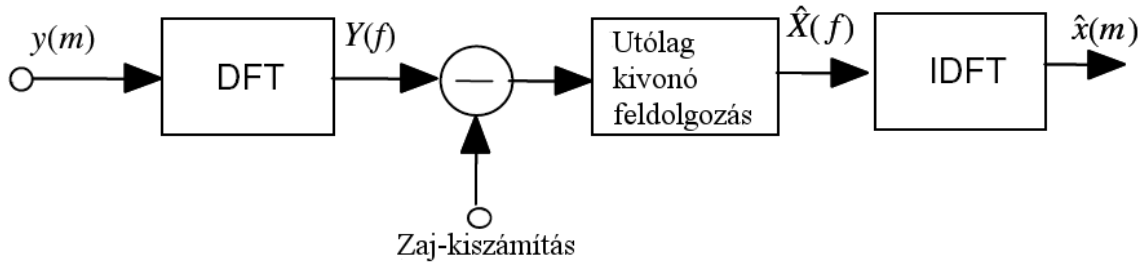


21. ábra Zajjal terhelt jelek alakjai az idő függvényében

zajos jel spektrumának középvértékéből kivonjuk a zaj spektrumának középvértékét. A zajos jel időben a következőképpen modellezhető:  $y(m)=x(m)+n(m)$ , ahol  $y(m)$  jelöli a zajjal terhelt jelet,  $x(m)$  a hasznos jelet,  $n(m)$  a zajt,  $m$  pedig a diszkrét idő mérőszám. Egy frekvenciatartományban az előző egyenlet a következőképpen fejthető ki:  $Y(f)=X(f)+N(f)$ , ahol  $Y(f)$  az  $y(m)$  zajos jel Fourier transzformáltja,  $X(f)$  az  $x(m)$  hasznos jel Fourier transzformáltja,  $N(f)$  az  $n(m)$  zaj Fourier transzformáltja, és  $f$  a frekvenciaváltozó. A spektrális kivonásban az  $y(m)$  zajos jelet vesszük, ezt egy átmeneti tárolóba helyezzük, majd felosztjuk  $N$  egyenlő részre. Minden szegmens esetében Hann vagy Hamming ablak segítségével szűrjük a spektrumot, ezután transzformáljuk diszkrét Fourier transzformációval (DFT) az összes  $N$  darab spektrális mintát. Az ablakok minden szegmens végén csökkentik annak hatását,

hogy a vett jelet feldaraboltuk, vagyis nem folytonos jeleket illesztünk össze. Az ablakozott jel így adható meg:  $y_w(m) = w(m)y(m) = w(m)[x(m) + n(m)] = x_w(m) + n_w(m)$ . Az alakozási eljárás a frekvenciatartományokra az alábbi módon fejthető ki (DFT segítségével):  $Y_w(f) = W(f) * Y(f) = X_w(f) + N_w(f)$ , ahol a  $*$  operátor a konvolúciós szorzást jelenti.

A következő ábra a spektrális kivonás módszerének blokkdiagramját szemlélteti. (Az IDFT jelölés a diszkrét Fourier transzformáció inverzét jelenti)



22. ábra A spektrális kivonás egyszerűsített blokkdiagramja

A spektrális kivonás egyenlettel a következő módon írható le:  $|\hat{X}(f)|^b = |Y(f)|^b - \alpha \overline{|N(f)|^b}$ , ahol  $|\hat{X}(f)|^b$  az eredeti jel spektrumának kiszámítása,  $|X(f)|^b$  és  $|N(f)|^b$  pedig az időátlagolt zaj spektrum. A spektrális kivonás nagyságához a  $b$  kitevő értékét 1-re, az erősségéhez pedig 2-re kell választani. Az  $\alpha$  paraméter a kivont zaj mennyiségét szabályozza. A teljes kivonáshoz  $\alpha=1$  értéket kell választani, a túlvonáshoz pedig  $\alpha>1$  érték lesz a megfelelő. Az időátlagolt zajspektrum beszerezhető a jelmentes szakaszokból, ahol csak kizárólag a zaj van jelen, nincs hasznos jel, és az alábbi összefüggés segítségével formalizálható:

$$\overline{|N(f)|^b} = \frac{1}{K} \sum_{i=0}^{K-1} |N_i(f)|^b, \text{ ahol } |N_i(f)|^b \text{ a spektrum } i\text{-edik zajkerete, és feltesszük,}$$

hogy a zaj  $K$  keretből áll, és  $K$  változó. Más esetben az átlagolt zajspektrum beszerezhető az elsőrendű digitális alul-áteresztő szűrő kimeneteként a következőképpen:

$$\overline{|N_i(f)|^b} = \rho \overline{|N_{i-1}(f)|^b} + (1 - \rho) |N_i(f)|^b, \text{ ahol az alul-áteresztő szűrő együtthatója}$$

$\rho$ , melynek tipikus értéke 0,85 és 0,99 közé szokott esni. Az időtartomány alapú jelvisszaállításhoz a nagyságspektrum  $|\hat{X}(f)|$  kiszámítását kombináljuk a zajos jel fázisával, majd ezt követően transzformáljuk időtartományba az inverz diszkrét Fourier transzformáció segítségével, mely a következőképpen írható fel:

$$\hat{x}(m) = \sum_{k=0}^{N-1} |\hat{X}(k)| e^{j\theta_Y(k)} e^{-j\frac{2\pi}{N}km}, \text{ ahol } \theta_Y(k) \text{ a zajos jel fázisa } Y(k) \text{ frekvencián. Ez a jel-}$$

viSSzaállító egyenlet feltételezi, hogy a hallható tartományba eső zaj főként a nagyságspektrum torzítása és a fázistorzítás miatt nagyrészt nem hallható. A szimulált fázistorzítás megfigyelhető hatásának becslése igazolta a feltevést. A zaj spektrumának változatossága következtében előfordulhat, hogy nagyság és az energiaspektrum szempontjából negatív eredményt kapunk a spektrális kivonással. Ennek végeredményeképpen nagyon valószínű, hogy csökkenni fog a jel-zaj arány. A negatív eredmények elkerülése érdekében a spektrális kivonás kimeneti jelét egy utófeldolgozó térképező függvénnyel feldolgozzuk, melyet a következőképpen formalizálhatunk:  $T[|\hat{X}(f)|] = \begin{cases} |\hat{X}(f)| & \text{ha } |\hat{X}(f)| > \beta |Y(f)| \\ \text{fn}[|Y(f)|] & \text{egyébként} \end{cases}$ . Például választhatunk egy

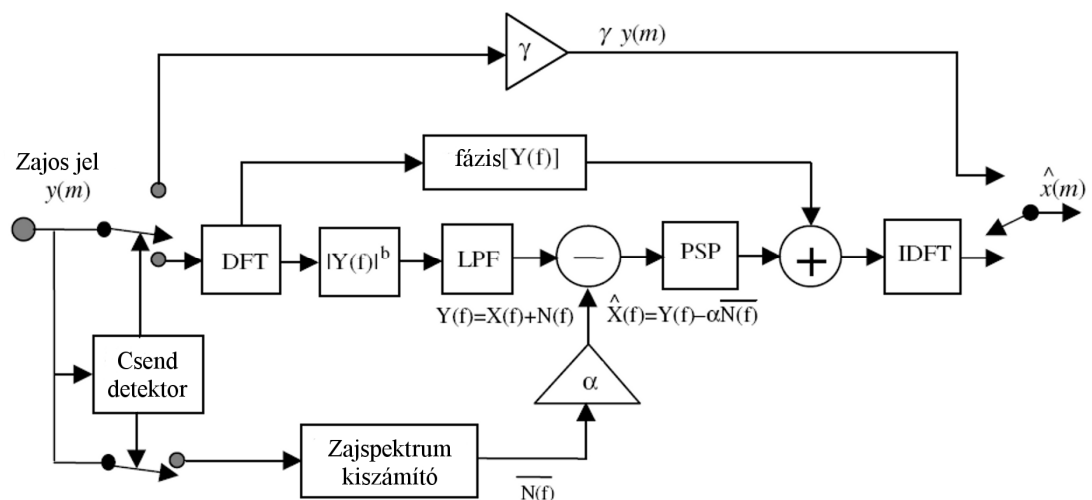
olyan szabályt, mely ha a  $|\hat{X}(f)| > 0,01 |Y(f)|$  (a nagyság spektrum szorzása 0,01-dal ekvivalens a  $-40\text{dB}$ -lel) érték igaz, akkor az  $|\hat{X}(f)|$  érték valamilyen függvénye legyen a zajos jelnek  $\text{fn}[|Y(f)|]$ . Ennek legegyszerűbb formája, hogy beállítunk egy  $\text{fn}[|Y(f)|] = \text{zajszint}$  értéket, ahol a zajszint pozitív konstans. Egy másik lehetőség, ha  $\text{fn}[|Y(f)|] = \beta |Y(f)|$ . Ez esetben:

$$T[|\hat{X}(f)|] = \begin{cases} |\hat{X}(f)| & \text{ha } |\hat{X}(f)| > \beta |Y(f)| \\ \beta |Y(f)| & \text{egyébként} \end{cases}$$

A spektrális kivonás energiaspektrum vagy nagyságspektrum tartományok alapján is implementálható. A két módszer hasonló, de elméletileg eredményeznek némi különbséget a várható teljesítményben.

### 3.3.3.2. A spektrális kivonó megvalósítása

Egy spektrális zajkivonó rendszer blokkdiagramját mutatja az alábbi ábra,



23. ábra Spektrális kivonó megvalósításának blokkdiagramja

mely a következő alrendszeréből áll:

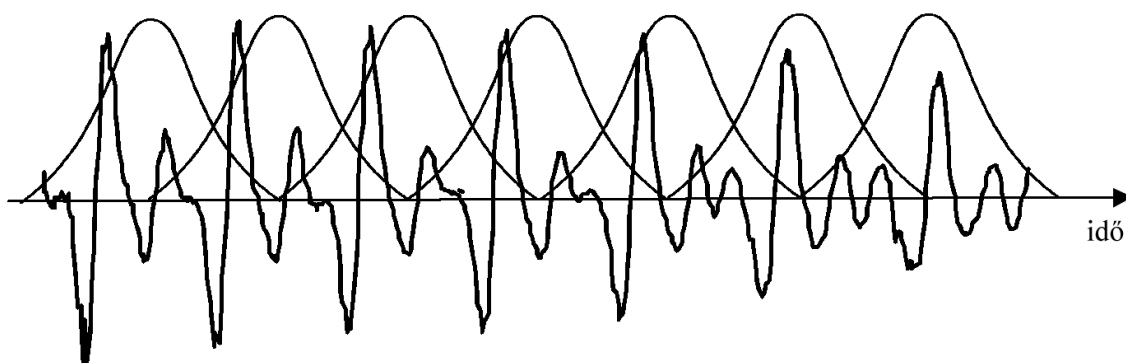
- a.) egy csend detektor a jelmentes időszakok megtalálására, a zaj spektruma ezen időtartamok alatt frissül.
- b.) a diszkrét Fourier transzformációt megvalósító rész (DFT) az időtartományt frekvenciatartományá alakítja, a DFT egységet egy nagyság operátor követi.
- c.) alul-áteresztő szűrő (Lowpass Filter, LPF), a zaj egyenetlenségeinek csökkentésére, melynek célja a zaj ingadozásaiból adódó feldolgozási torzítás csökkentése.
- d.) utófeldolgozó, a spektrális kivonás eredményeképpen keletkezett torzítások eltávolítására.
- e.) inverz diszkrét Fourier transzformációt végző rész, mely frekvenciatartományból viszi a feldolgozott jelet időtartományba.
- f.) csillapító berendezés  $\gamma$ , mely a zaj csillapítását végzi a csöndes részeken.

A diszkrét Fourier transzformáción alapuló spektrális kivonó blokkdiagramja egy algoritmus. A beérkező audiojelet átmeneti tárolóba rakjuk, majd  $N$  darab egymást átfedő mintablokkra osztjuk. Minden blokk Hann vagy Hamming ablakolt, és azt követően diszkrét Fourier transzformációval átalakítjuk frekvenciatartományá. A spektrális kivonás után a jel nagyságspektrumát kombináljuk a zajos jel fázisával, és visszaalakítjuk időtartományá. Minden blokkot egymáshoz illesztünk, úgy, hogy legyen átfedés a megelőző és a rákövetkező blokkal, és ez adja a végleges kimeneti jelalakot.

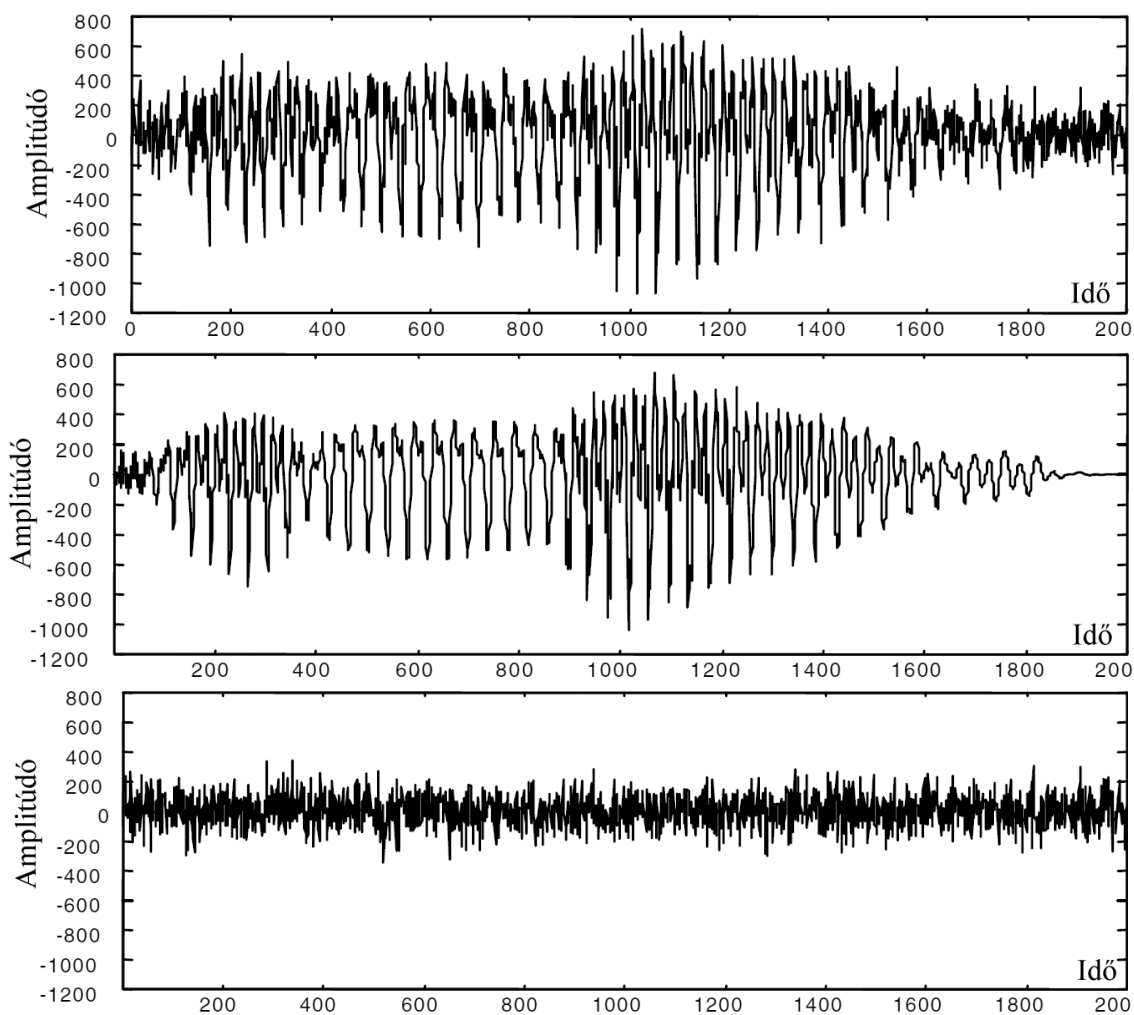
A blokk hosszának megválasztása a spektrális elemzéshez egy kompromiszsum, mely két ellentmondó szükséglet, az időfelbontás és a spektrális felbontás igényeit kell, hogy kielégítse. Tipikusan 5–50 ms közé esik a használt blokkok hossza. Ez egy 20 kHz-es mintavételezési frekvencián azt jelenti, hogy  $N$  értéke 100–1000 tartományba esik. A spektrum frekvenciafelbontása egyenesen arányos a minták számával ( $N$ ). Nagyobb  $N$  esetén a spektrum értékelése jobb lesz. Ez csak részben igaz a frekvenciatartomány alsó részére, mivel az alacsonyabb frekvenciájú részek az idő múlásával lassan változnak, és nagyobb ablak szükséges a stabil értékeléshez. Az ellentmondó szükséglet az audiojelek nem stacionárius természetéből adódnak, tehát az ablak méretének azért nem szabad túl nagyoknak lennie, hogy a rövid időtartamú események ne mosódjanak el.

Az ablakok átfedéseinek (24. ábra) fő feladata, hogy csillapítsák a blokkok végpontjain keletkező folytonossághiányokat, melyek a darabolásból adódnak. Bár



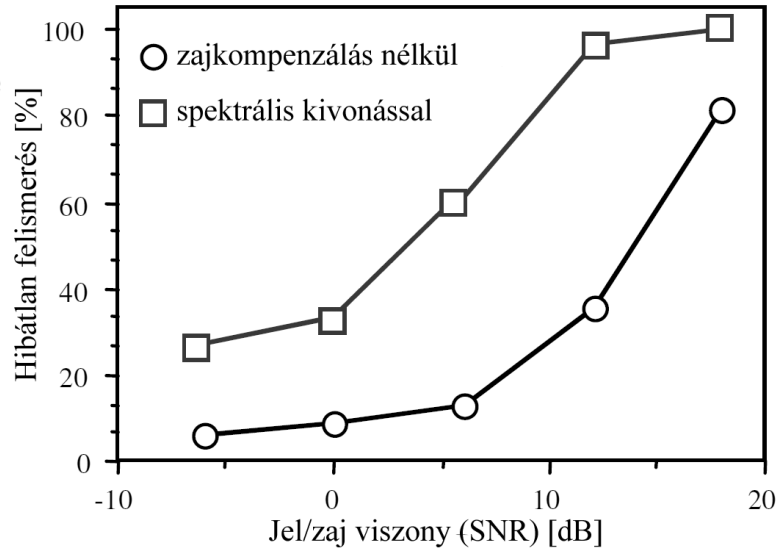


24. ábra A spektrális kivonásban alkalmazott ablakolási és átfedési folyamat létezik számos felhasználható ablaktípus, különböző karakterisztikákkal, mégis a legtöbb megvalósításban Hamming ablakokat használnak. A torzítás eltávolítás során – melyet a spektrális kivonás eredményezett – az utófeldolgozó algoritmus olyan információkat használ fel, mint a frekvenciacsatornák összefüggése az egymást követő blokkokban, valamint az jel eseményeinek és torzításának időtartamai.



25. ábra Zajos jel (felső), zajmentes jel (középső), és zaj (alsó)

A jel spektrális összetevőinek összefüggése – az időtengely mentén – részben szabályozható az ablakhossz megválasztásával és az átfedéssel. Ez az összefüggés növekszik az ablak hosszának csökkentésével és az átfedés növelésével. Az átfedés növelése azonban növeli a zajfrekvenciák összefüggését is, az időtengely mentén.



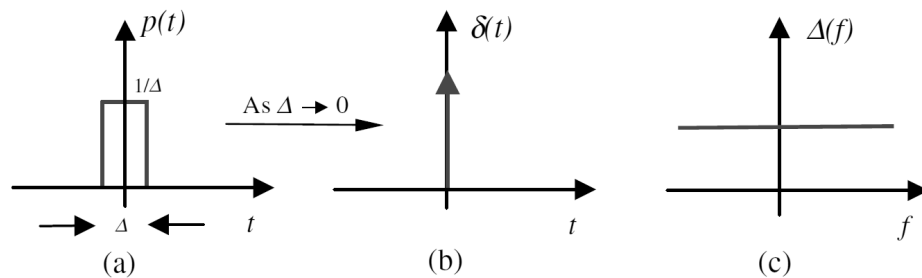
26. ábra A spektrális kivonás jelentősége a beszéd felismerésben

### 3.3.4. Az impulzív zaj eltávolítása

Az impulzív zajok gyors be/ki kapcsolásokhoz hasonló impulzusokat produkálnak és nagyon kiszámíthatatlan a megjelenésük, tehát bármely időpillanatban előfordulhatnak, és bármekkora időközrel. Ezen zajok eltávolítására és csökkentésére az előző részben kifejtett spektrális kivonás módszere nem alkalmas, mert itt a zajszint nem határozható meg.

#### 3.3.4.1. Impulzív zaj matematikai modellje

Az analóg impulzív zajt matematikai értelemben az alábbi ábra segítségével szemléltethetjük. Ha megvizsgáljuk az ábra (a) részén szemléltetett egységnyi te-



27. ábra (a) Egységnyi területű lökés, (b) impulzus, (c) az impulzus spektruma

rületű lökeshullám  $p(t)$  (négyzög jel) viselkedését, ahogy a lökés szélessége  $\Delta$  nullához tart, a lökés impulzussá válik. Az impulzus függvényt az ábra (b) része mutatja, és egy végtelenül kicsi időszélességgel rendelkező lökést a következőképpen

definiálhatunk:  $\delta(t) = \lim_{\Delta \rightarrow 0} p(t) = \begin{cases} \frac{1}{\Delta}, & |t| \leq \frac{\Delta}{2} \\ 0, & |t| > \frac{\Delta}{2} \end{cases}$ . Az impulzusfüggvény integrálja a kö-

vetkező:  $\int_{-\infty}^{\infty} \delta(t) dt = \Delta \times \frac{1}{\Delta} = 1$ . Az impulzusfüggvény Fourier transzformáltja felállítha-

tó:  $\Delta(f) = \int \delta(t) e^{-j2\pi ft} dt = e^0 = 1$ , ahol  $f$  a frekvenciaváltozó. Az impulzusfüggvényt

teszt jelleggel is szokták alkalmazni, a rendszer impulzusreakciójának megállapítására. Az ábra (c) részén mutatottak szerint az impulzus egy spektrálisan nagyon gazdag jel, minden frekvenciát azonos mennyiségben tartalmaz, tehát ezért alkalmas a tesztelésre.

Egy digitális impulzus  $\delta(m)$  (8. oldal 4. ábra (a) részén látható) úgy definiálható, mint egy egy minta hosszúságú „be” jel, és kifejezhető:  $\delta(m) = \begin{cases} 1, & m = 0 \\ 0, & m \neq 0 \end{cases}$ , ahol az

$m$  változó jelöli a diszkrét idő mérőszámot. A Fourier transzformációt használva

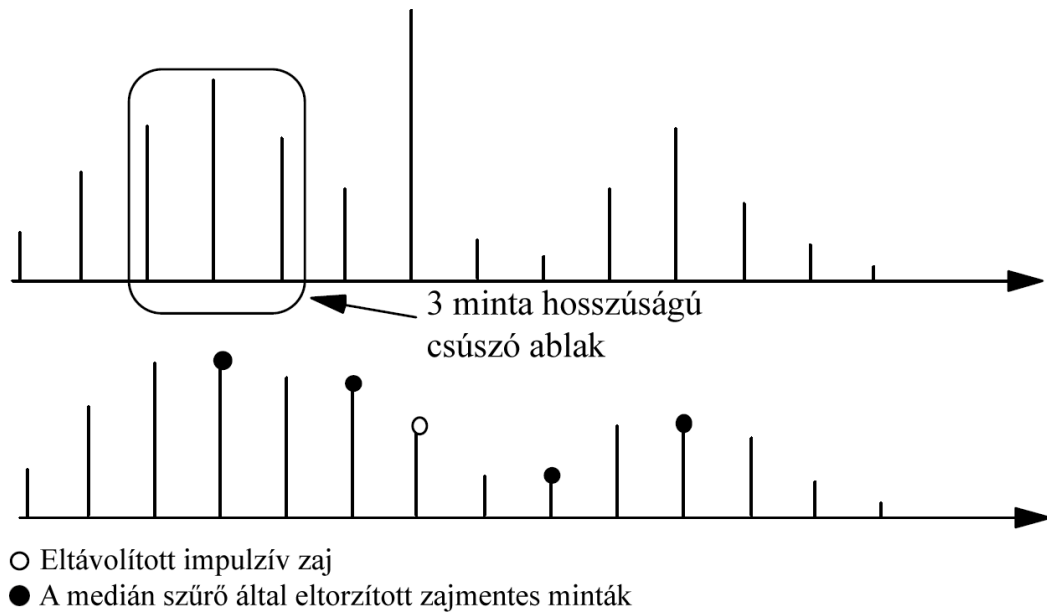
egy digitális impulzus frekvenciaspektruma megadható:  $\Delta(f) = \sum_{m=-\infty}^{\infty} \delta(m) e^{-j2\pi fm} = 1, 0,$

$-\infty < f < \infty$ .

Kommunikációs rendszerekbe a valódi impulzív típusú zajok hossza normálisan nagyobb, mint egy minta.

#### 3.3.4.2. A medián szűrő

Az impulzív zaj eltávolításának klasszikus megközelítése a medián szűrő. Medián alatt értjük az  $\{x(m)\}$  mintahalmaz elemeiből azt az  $x_{\text{med}}(m)$  értéket, melyre igaz, hogy a halmaz elemeinek fele ezen  $x_{\text{med}}(m)$  értéknél nagyobb, a másik fele pedig az  $x_{\text{med}}(m)$  értéknél kisebb. Tekintsük a mediánt minták egy halmazának, melyet úgy kapunk meg, hogy a mintákat rendezzük növekvő vagy csökkenő sorrendbe és kiválasztjuk a középső értéket. A medián szűrés esetén egy előre meghatározott hosszúságú ablakot csúsztatunk folyamatosan a jelen, és az ablakban található középső mintát kicseréljük az ablak mintáinak mediánjára, ahogy azt a következő ábra mutatja.



28. ábra Medián szűrő be- és kimenete. A működés közben torzítás is keletkezhet

A medián szűrő  $\hat{x}(m)$  kimenete az  $y(m)$  bemenet és egy  $2K+1$  mintát tartalmazó hosszúságú medián ablak alapján:

$$\hat{x}(m) = y_{\text{med}}(m) = \text{medián}[y(m-K), \dots, y(m), \dots, y(m+K)]$$

Egy számhalmaz mediánja a halmaz nem lineáris statisztikája egy nagyon jól hasznosítható tulajdonsággal, hogy nem érzékeny egy szokatlanul nagy mintaérték jelenlétére a halmazban, melyet kívülállónak hívunk. Ezzel szemben a középérték és részben a variancia érzékeny egy számhalmazban jelen levő kiugró értékre, vagyis az impulzív típusú zajra. Fontos tulajdonsága a medián szűrőnek – mely részben használható a képfeldolgozásban –, hogy megőrzi az éleket vagy a lépésenkénti folytonossághiányokat a jelben. A medián szűrők hatékonyan használhatók a képfeldolgozásban az impulzív zaj eltüntetésére az élek elmosása nélkül – ez nagyon jelentős a képfeldolgozásban. A medián filterrel történő kísérletezés során kiderült, hogy nem alkalmas magas minőségű eredmények produkálására audiojelek visszaállítása esetében. A medián szűrő nem tud mit kezdeni az olyan impulzív jellegű zajjal, mely hosszabb, mint egy vagy két minta. Továbbá a medián szűrő a hibásan impulzív zajnak vélt jelek módosításával jelentős mennyiségű feldolgozási torzítást okoz. A medián szűrő teljesítménye javítható egy alkalmazkodó határérték bevezetésével, tehát a mintát csak akkor cseréljük ki, ha annak eltérése a középső érték-

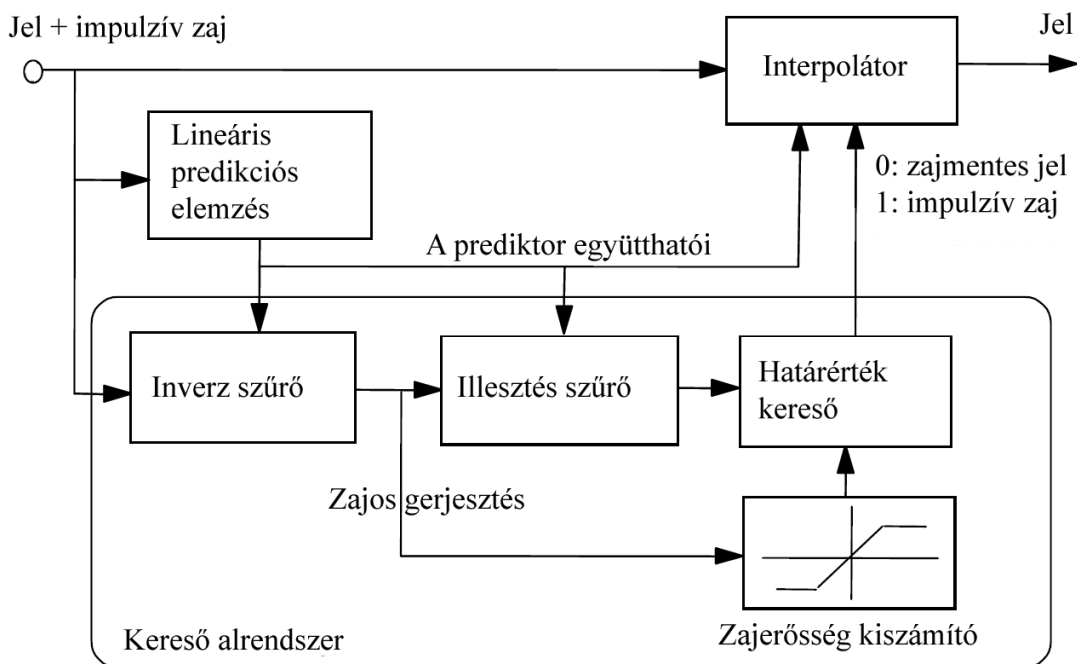
től meghaladja a határértéket:  $\hat{x}(m) = \begin{cases} y(m) & \text{ha } |y(m) - y_{\text{med}}(m)| < k\theta(m) \\ y_{\text{med}}(m) & \text{egyébként} \end{cases}$ , ahol

$\theta(m)$  egy alkalmazkodó határérték, mely erőteljes becsléssel határozza meg

$|y(m) - y_{\text{med}}(m)|$  átlagértékét,  $k$  pedig egy hangoló paraméter. A medián szűrők nem optimálisak, mert nem használhatók hatékonyan a jelgenerálás fiziológiájának előismerete vagy a jel és a zaj statisztikai eloszlásának modellezése nélkül.

### 3.3.4.3. Impulzív zaj eltávolítása lineáris prediktív modell segítségével

Az impulzív zavarok általában a jel egészét vizsgálva csupán csekély  $\alpha$  hányadát szennyezik. Mivel a jel jelentős  $1-\alpha$  része impulzív zajjal nem szennyezett, ezért előnyösebb megkeresni a zajos részeket (egyedülálló impulzusokat), és csak a torzult mintákat javítani. Ez a stratégia kiküszöböli a felesleges feldolgozásokat, és a minták egy viszonylag nagy része – mely nem szennyezett impulzív zajjal – érintetlenül eredeti állapotában marad. A következő ábra egy ilyen rendszer blokkdiagramját mutatja, mely két alrendszerből áll, egy keresőből és egy interpolátorból. A



29. ábra Egy impulzív zajeltávolító blokkdiagramja.

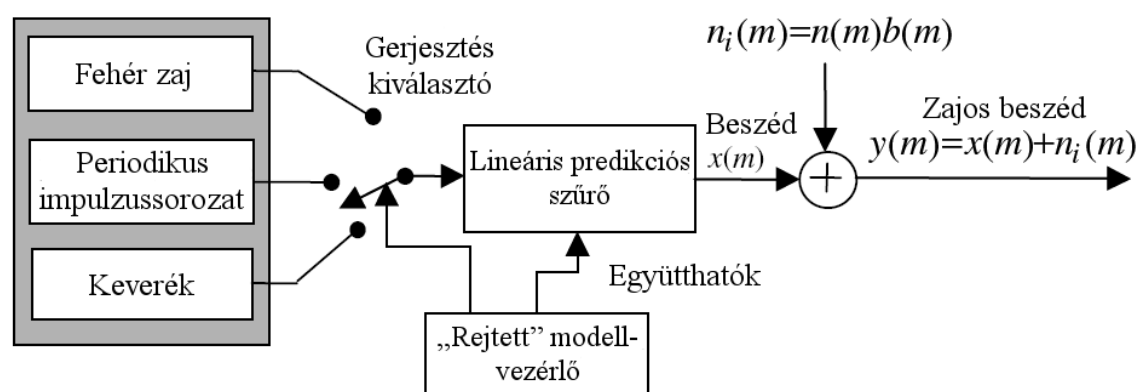
detektor megkeresi az impulzusok helyét, az interpolátor pedig kicseréli a torzult mintákat, figyelembe véve a torzult minta mindkét oldalán elhelyezkedő további mintákat. A detektor rész egy lineáris predikciós elemzőből, egy illeszkedésvizsgáló szűrőből és egy határérték keresőből áll. A kereső kimenete egy kétállású kapcsoló, mely az interpolátort vezérli. Impulzív zaj hiányában az interpolátor nincs bekapcsolva. Impulzív zaj jelenléte esetében az interpolátor bekapcsol és kicseréli a zaj által törölt mintákat.

## 3.3.4.3.1. Az impulzív zaj keresése

A legegyszerűbb módszer az ilyen jellegű zajok megtalálására, hogy meghatározzunk egy határértéket, és a határérték fölé eső mintákat zajnak osztályozzuk. Ez a módszer tökéletesen működik nagyon nagy impulzusok esetében, de ha a zaj amplitúdója a megadott határérték alá esik, akkor nem találja meg. A keresés eredményesebbé tehető, ha kihasználjuk a jel és az impulzív zaj jellemzőinek különbségét. Az impulzív zaj, vagy rövid időtartamú hullámvész bevisz egy a jelre nem jellemző megszakítást a jelbe. A megszakítás könnyebben megtalálható, ha a jelet szétválasztjuk. A szétválasztás (vagy digitális jelnél megkülönböztetés) művelete egyenértékű a kölcsönösségi viszony megszüntetésével, vagy a spektrális fehérítéssel. A lineáris prediktor modellezi a jel kölcsönösségi struktúráját, és ezen összefüggések megszüntetését végzi az inverz szűrő. A zajmentes  $x(m)$  jel lineáris predikciós modell segítségével a következő összefüggéssel írható fel:

$$x(m) = \sum_{k=1}^P a_k x(m-k) + e(m), \text{ ahol } a=[a_1, a_2, \dots, a_P]^T \text{ a lineáris prediktor } P \text{ rendű}$$

együttható vektora, és az  $e(m)$  gerjesztés egy zajhoz hasonló jel, vagy vegyes véletlenszerű zaj keveréke, és hullámvészok kvázi periodikus sorozata, melyet a következő ábra mutat. Az impulzív zajkereső nem a korrelálatlan kétállapotú impulzív zaj



30. ábra Zajos beszédmodell. A jelet egy lineáris prediktor modellezi, az impulzív zaj modellezésére pedig egy kétállapotú amplitúdómodulált modellt használunk.

alapján dolgozik, hanem azon a megfigyelésen alapszik, hogy a lineáris prediktorok jó modellt biztosítanak korrelált jelekre. A zajos  $y(m)$  jel a prediktor gerjesztett jelévé történő alakítása a következő hatásokat váltja ki:

- a.) Az jel amplitúdójának léptéke csökken az eredetileg gerjesztett jeléhez képest, mivel a zaj amplitúdójának léptéke változatlan marad, vagy növekszik.

- b.) A jel nem lesz kölcsönösen összefüggő, mert az impulzív zaj elmaszatólódott és átalakult léptékes alakba az inverz szűrő impulzus válaszaként.

Mindkét hatás javítja az impulzív zaj megtalálhatóságát. A beszéd és a zene hang időtartamának rezonanciája vagy a hangszer által erősített és spektrálisan alakított gerjesztésekből áll össze. Ez a gerjesztés véletlenszerűbb a beszédnél, és gyakran sokkal kisebb az amplitúdó-tartománya. Az impulzív zaj megtalálhatósága javítható az inverz szűréssel, mely lényeges lehet és függ a jel időben változó korrelációs struktúrájától. Ez a módszer egy optimális határértékképző használatával képes hatékonyan csökkenteni az impulzív zaj keresésének problémáját, a sorból kilógó minták megtalálását.

#### 3.3.4.3.2. Az impulzív zaj megtalálhatóság-javításának elemzése

A következőkben az impulzív zaj megtalálhatóságának javítását elemezzük ez inverz szűrő kimenetének vizsgálatával. Az előző egyenlet használatával a zajos jel modelljét a következő egyenlet segítségével (az előző átírt változata) írhatjuk fel:

$$y(m) = x(m) + n_i(m) = \sum_{k=1}^P a_k x(m-k) + e(m) + n_i(m), \text{ ahol } x(m), y(m) \text{ és } n_i(m) \text{ rendre a}$$

zajos jel, a jel és a zaj. Egy számítást használva  $\hat{\mathbf{a}}$  lesz  $\mathbf{a}$  prediktor együttható vektora, és a zajos  $y(m)$  jel inverz szűrhető és átalakítható  $v(m)$  zajosan gerjesztett

$$\text{jellé így: } v(m) = y(m) - \sum_{k=1}^P \hat{a}_k y(m-k) = x(m) - n_i(m) - \sum_{k=1}^P (a_k - \tilde{a}_k)[x(m-k) + n_i(m-k)],$$

ahol  $\tilde{a}_k$  a prediktor együttható számítási hibája. A zajos jel modelljének egyenlete ez alapján az egyenlet alapján következő egyenleté írható át egy másik alakba:

$$v(m) = e(m) + n_i(m) + \sum_{k=1}^P \tilde{a}_k x(m-k) - \sum_{k=1}^P \hat{a}_k n_i(m-k). \text{ Ebből az egyenletből három}$$

tétel határozható meg, mely elősegíti a zaj megtalálását a gerjesztési szekvenciában.

- a.) az  $n_i(m)$  impulzív zavar, mely általában domináns;
- b.) a régebbi  $P$  darab impulzív zajminta hatása az inverz szűrő működése végett, a jelenbe maszátolva:  $\sum \hat{a}_k n_i(m-k)$ ;
- c.) a gerjesztett jel varianciájának növekedése a paramétervektor számítási hibái miatt, mely a következő kifejezéssel írható le:  $\sum \tilde{a}_k x(m-k)$

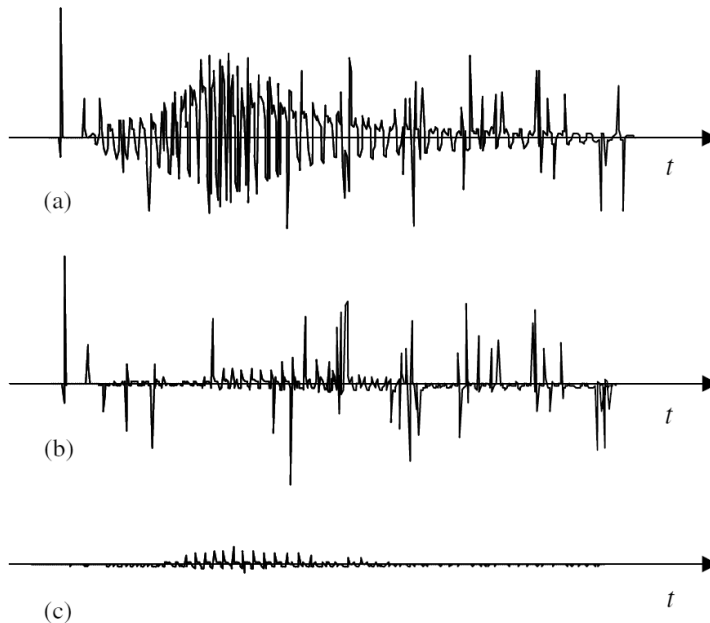
Az inverz szűrő eredményeinek javítása a következő formulákkal írható le. Az impulzív zaj a zajos jelhez viszonyított aránya az alábbiakban ismertetett módon

adható meg:  $\frac{\text{impulzív zaj energiája}}{\text{jel energiája}} = \frac{E[n_i^2(m)]}{E[x^2(m)]}$ , ahol  $E[\cdot]$  a várható érték operátora.

Az impulzív zaj detektálásában az érdeklődés tárgyát képező jel az impulzív zaj lesz, melyet a kísérő jelből kell megtalálni. Feltesszük, hogy a domináns zaj a  $v(m)$  zajos gerjesztési jelben az  $n_i(m)$  impulzív zaj, a két jel aránya a következőképpen írható fel:  $\frac{\text{impulzív zaj energiája}}{\text{gerjesztési energia}} = \frac{E[n_i^2(m)]}{E[e^2(m)]}$ . A teljes impulzív zaj jellel szembeni

aránya az előző két egyenlet hányadosából kapható meg, melyet a következő egyenlettel tudunk leírni:  $\frac{E[x^2(m)]}{E[e^2(m)]} = \text{nagyobbodás}$ .

Ez az egyszerű elemzés megmutatta, hogy az impulzív zaj megtalálhatósága függ a teljesítményerősítés karakterisztikájától, ennek következtében a lineáris predikciós modell rezonanciáitól. Beszéd esetében a jel amplitúdója zajmentes gerjesztett jelről lévén szó  $10^{-1}$  és  $10^{-4}$  között változik, ami a beszéd maga, ezért jelentős javulás érhető el az impulzív zaj megtalálhatóságában az inverz szűrő alkalmazásával a zajos beszédre. A következő ábra az impulzív zaj megtalálhatóságának



31. ábra Az inverz szűrő hatása a impulzív zaj megtalálhatóságára. (a) impulzív zajjal szennyezett beszéd, (b) az impulzív zajjal szennyezett beszéd gerjesztése, (c) beszéd gerjesztés impulzusmentesítve.

gának javítására alkalmazott inverz szűrő hatását mutatja. Az inverz szűrés hatása, hogy az  $x(m)$  jelet korrelálatlan  $e(m)$  jelgerjesztéssé alakítja, míg az impulzív zaj az inverz szűrő impulzív  $[1, -a_1, \dots, -a_p]$  zajátvitelének megfelelően torzul és megnövekszik, ahogy azt az iménti tétel mutatja a következő összefüggés alapján:

$\sum \hat{a}_k n_i(m-k)$ . Feltételezük, hogy a gerjesztés egy

normális eloszlást követő fehér zaj, egy szűrő, melynek együtthatói az inverz szűrő együtthatóival egyeznek képes növelni a torzult impulzív zaj megtalálhatóságát a gerjesztési jelben.



## 3.3.4.3.3. Kétoldali prediktor az impulzív zaj megtalálására

Az impulzív zaj megtalálásának egyik hatékony módja, hogy a jelet korrelálatlanná tesszük. Ez a módszer felhasználható a kétoldali lineáris predikciós modellben is. Egy  $x(m)$  minta kétoldali lineáris predikciója  $P$  darab régebbi (múltbeli) és  $P$  darab jövőbeli minta alapján dolgozik, és a következő egyenlettel definiálható:

$$x(m) = \sum_{k=1}^P a_k x(m-k) + \sum_{k=1}^P a_{k+P} x(m+k) + e(m),$$

ahol  $a_k$  értékek a kétoldali prediktor együtthatói, és  $e(m)$  a gerjesztési jel. Minden elemzés arra az esetre készül, mikor csak egy oldalról vizsgáljuk a jelet, de ez kiterjeszthető a kétoldali modellre is. Azonban a gerjesztés mértéke a kétoldali modell esetében kisebb varianciát mutat, mint az egyoldali modellben, mivel az imént leírt egyenletben mind a múltbeli és mind a jövőbeli összes mintára modellezett a minták korrelációja. Habár az iménti egyenletben egy nem véletlenszerű szűrő van, az inverze szükséges a kereső alrendszerbe, mely már véletlenszerű. A kétoldali prediktor használata további javulást eredményezhet az impulzív zaj megtalálhatóságában.

## 3.3.4.3.4. Az eldobott minták interpolációja

Az impulzív zaj hatására a minták visszavonhatatlanul torzulnak, és eltávolításukkal „űr” keletkezik, melyet interpolációval próbálhatunk kitölteni. A hiányosságok interpolálására – ahhoz, hogy ne maradjon hallható nyomuk –, magas hanghűségű interpolátor szükséges. A legkisebb négyzetes autoregresszív (LSAR) módszer magas minőséget eredményez viszonylag kis számú az impulzív zaj miatt elveszett minta esetében. Az LSAR interpolációs módszer egy kétszintű folyamat. Az első szinten az impulzív zaj mindkét oldalán megmaradt minták alapján becsléssel meghatározzuk a jel predikciós modelljének paramétereit. A második szinten a meghatározott modellparaméterek és a hézag két oldalán található minták alapján interpoláljuk az elveszett mintákat. Ezzel az interpolációs módszerrel az impulzív zaj által eltorzított audiojelek magas minőségben visszaállíthatók.

## 4. Zajsűrés a gyakorlatban

### 4.1. Zenekari stúdiómunka

#### 4.1.1. A felvétel

Egy hangstúdióban legyen szó akár könnyűzenei vagy komolyzenei műről a hangrögzítés során számos problémával kell szembenéznünk. Amennyiben megoldható, az akusztikus (nem elektromos) hangszerek felvételét úgynevezett süket-szobában célszerű kivitelezni, a visszhanghatások elkerülése végett, melyek a hanganyag torzulását okozhatják. A felvételt adott esetben csatornákra bontva külön-külön rögzítjük. Ez azt jelenti, hogy az egyes hangszerek külön játsszák a darabot, majd a felvételt követően egy keverőpult segítségével születik meg a végleges hanganyag.

A felvétel során is felütik a fejüket a zajok. Amivel mindenképpen számolni kell, hogy a bemikrofonozott jelforrás valamilyen távolságra helyezkedik el a mikrofontól, néha a zenész átéléssel játszik, és a hangszerrel együtt mozog. A távolság és az alkalmazott mikrofon minőségének függvényében a rögzítő technika mindenképpen rögzít egy kevés környezeti zajt is. Ezen zajok ellen a környezet megfelelő megválasztásával lehet a legkönnyebben védekezni.

További problémákat okoznak a mikrofonok viszonylag alacsony jelszintjei, és a jeleket továbbító kábelek. A kábeleket célszerű lehetőség szerint minél vastagabbra, tiszta anyagból készültre, és minél rövidebbre választani, a lehető legkevesebb csatlakozóponttal, ugyanis ez tudja a jel legjobb minőségű szállítását elvégezni. A nagy keresztmetszet és a viszonylag kis hossz biztosítja a kis ellenállást, mely a jelveszteséget csökkenti, és a tiszta anyag (például oxigénmentes réz, négy kilences tisztasággal – 99,99 % réz) pedig a szennyezések miatt kialakuló örvényáramok okozta zajtól óvnak bennünket. A csatlakozópontok további torzítást és zajt eredményezhetnek, az anyagok különbözősége, és a csatlakozó felületek mérete miatt. További eredményjavulás várható az arany, vagy felületén aranyozott kábelek alkalmazásától a magas frekvenciájú jelek átvitele szempontjából – a skin effektus miatt. A skin effektus azt jelenti, hogy a váltakozó áramot szállító vezetékben az elemi töltéshordozók a vezetőben annál távolabb haladnak a vezető középpontjától, minél nagyobb az átvitt jel frekvenciája.

Az alacsony jelszint miatt a digitális feldolgozhatóság elérése érdekében előerősítő fokozatokat kell a rögzítő készülék és a jelforrás(ok) közé építeni, melyek esetében az alkatrészek kialakítástól és minőségétől függően termikus zajgenerátorként működnek. A sávátviteli problémákat többnyire nem az erősítőelemek okozzák, hanem a mikrofonok.

További érdekesség a jelforrás szempontjából, hogy az elektromos hangszereket, mint például egy elektromos gitár, vagy basszusgitár általában egy úgynevezett alap erősítőre csatlakoztatjuk, majd ezen erősítő hangszórója elé helyezünk el mikrofont. Általában e hangszerek esetében az effekteket már az erősítő előtt, vagy annak effektútjába csatlakoztatjuk. Ilyenkor egy torzító effekt képes a hangszer hangszedője (pickupja) által felfogott zajokat nagymértékben felerősíteni, amit még – egy a torzító előtt elhelyezett (a játéktechnikát könnyítő) – dinamika kompresszorral lehet „növelni”. Ezek a zajok elektromágneses zajok, és sajnos a hangszer működéséből adódóan nehéz ellenük védekezni, mivel a húrok rezgését is ilyen módon érzékeli a hangszedő. Ilyen hangszerek esetén már két termikus zajgenerátor rendszer kapcsolódik a rögzítő elé, ezzel fokozva a zajszennyezést.

A rögzítő technika manapság már többnyire digitális rendszer, melynek az érkező jeleket egy ADC segítségével digitálissá kell alakítania. A jel ekkorra már zajjal erősen szennyezett lesz, a korábbi berendezések felépítésének és minőségének függvényében. A digitalizálás a mai korszerű hangkártyáknak nem jelentenek problémát, mégis stúdió célokra külön gyártanak – magasabb árfekvésű, jobb minőségű – kártyákat. A mintavételezési frekvencia megválasztása ekkor az egyik fő probléma. Általában a végeredmény CD lemezre kerül, ahol szabvány írja elő a 44100 Hz-es mintavételezési frekvenciát. Vannak hangkártyák, melyek ilyen mintavételezési frekvencia mellett nem produkálnak lineáris sávátvitelt, ellenben 48000 Hz-en igen. A felvett jelet visszahallgatva, van különbség 48000 Hz-en és a 44100 Hz-en mintavételezett jelek között az első javára, bár azt gondolhatnánk, hogy a hallható tartományon már ezek a frekvenciák kívül esnek. Másik érdekes kérdés a kvantálási hossz. A mai világban már a megfelelő dinamika elérése érdekében 16 bitnél nagyobb értéket választunk.

#### 4.1.2. Az utómunka

Az utómunkálatok első lépése a felvétel során a jelre került zajok eltávolítása. Ez esetben, ha sávonként rögzítettünk, akkor minden sávra külön célszerű a zajt eltávolítani, egészen pontosan annak hatását csökkenteni, ugyanis a zajok vélet-

lenszerű mivolta nem teszi lehetővé teljes eltávolításukat. Az ilyen felvételek esetén azért kell külön kezelni a zajokat, mert egy mikrofon és egy elektromágneses hangszedő több zajgenerátorral együttműködve más és más jellegű zajokat eredményeznek. Ha például egy dobfelszerelésre gondolunk, akkor észrevehető, hogy a megfelelő minőségű jel érdekében célszerű a hangszer egyes darabjait külön be-mikrofonozni, adott esetben más típusú mikrofonokkal, tehát ekkor már egy mikrofonrendszerrel beszélünk. Az egyes mikrofonok hangját külön kell bontani hangszín szempontjából, esetleg sztereó felvétel esetében még a panoráma sem fog egyezni minden esetben (például: lábcintányér jobboldal, pergődob közép, kísérő cintányér baloldal, természetesen a zenész kezességét figyelembe véve lehet fordítva is). A hangszínt minden esetben az egyes részeken külön kell beállítani, ami függ a mikrofon minőségétől, távolságától, helyzetétől és állásától. Ez a mikrofonrendszer mikrofononkénti hangszín beállításokkal másfajta zajt állít elő, mint például az imént említett elektromos gitár hangszedője egy torzítóval. Ezért is fontos a sávonkénti felvétel. Ezek a zajok csak ritka esetben tartalmaznak impulzív jellegű zajokat (vagy ha az ilyen zajok megjelennek, azok megszüntetéséről még a felvétel előtt célszerű gondoskodni), ezért a zaj csökkentésére a legalkalmasabb módszer a spektrális kivonás. Fontos, hogy a zajszűrő program (algoritmus) olyan legyen, hogy a hangképet ne, vagy csak nagyon kis mértékben módosítsa, a legjobb eredmény elérése érdekében.

Ezt követően az egyes hangszerek még további effektek hozzáadásával színesíthetők, valamint itt nyílik lehetőség az utolsó hangszínbeállításra, ugyanis a felvételkor még csak nyersanyag keletkezik. Itt adott esetben az egyes sávokat további analóg keverőpultba küldhetjük, de léteznek olyan stúdióban alkalmazható programok, melyek képesek a többsávós felvétel kezelésére. Ilyen program például a korábbi (a 2003-ban az Adobe cég által megvásárolt Syntrillium cég fejlesztésében megjelent) Cool Edit, és a manapság helyette alkalmazott Adobe Audition program, melyek kereskedelmi forgalomban kaphatók, de van ingyenes időkorlátos próbaváltozatuk is. Ezek a programok nem csak többsávós hanganyag effektezésére, hangszínbeállítására alkalmasak, hanem az azt megelőző zajcsökkentésekre is.

Természetesen a megfelelő hangzás elérése érdekében a hanganyagot a hangfelvétel elkészülte után hetekkel célszerű keverni, majd a végeredményt többféle berendezésen visszahallgatni, a maximális minőség elérése érdekében. A megfelelő hangzás elérése után még egy dinamika kompresszor segítségével lehet az anyagot

„összetömöríteni”. A mintavételezési frekvencia és a kvantálási hossz a későbbi hanghordozó függvényében változhat a rögzítési értékekhez képest, de ez sajnos a jel torzulását és néha még további zajok megjelenését is okozhatja.

## **4.2. Hagyományos hanglemez, hangkazetta archiválása CD-re**

Minden archiválási folyamat az adott hanghordozón tártolt hanganyag digitalizálásával kezdődik. Ezt akár otthonunkban is megtehetjük, feltéve, hogy rendelkezünk a forrás hanghordozó megfelelő minőségű lejátszására alkalmas berendezéssel. A hangminőséget jelentősen befolyásolhatja a jelforrás, továbbá az is, hogy milyen eszközzel végezzük a digitalizálást, és hogy milyen kábelek kötik össze az adott berendezéseket. Itt is lehet környezeti elektromágneses zavarokra számítani.

### *4.2.1. Hanglemez archiválása*

#### 4.2.1.1. A lejátszó rendszer

A gyártók általában többféle megoldást alkalmaznak a lemezejátszó mechanikájának és a mechanikai mozgásokat elektromos jellé alakító részegységek kivitelezésére.

A mechanika leglényegesebb elemei:

- Lemeztányér, amely a lemez forgatásáért, megtartásáért felelős, továbbá sok esetben a hangkar mozgatásával is foglalkozik, az alsó mechanikus szerkezeten keresztül. Többnyire nagy tömegű, az egyenletes forgás megvalósítása érdekében (nagyobb tehetetlenség, így nem képes a motor gyors fordulatszám-változásainak követésére). Fontos a jó csapágyazás és kenés a könnyű futás érdekében.
- Motor, amely a lemeztányért hajtja, nagyon fontos, hogy állandó fordulatszámú legyen nagyon kicsi ingadozással, mely valamilyen általában erőzárás közlőművel kapcsolódik a lemeztányérhoz (ez esetben fordulatszáma sokkal nagyobb, mint a tányéré).
- Hangkar (esetleg a mozgató mechanikával együtt), mely hangszedő fejjel együtt a lemeztűt egy shellnek nevezett tartóban rögzíti, és gondoskodik a megfelelő tűnyomásról és a barázdakövetésről.

A lemeztányér meghajtására háromféle megoldást szoktak alkalmazni, az egyik a dörzshajtás, a másik a szíjhajtás – gyakorlatilag ez az egyik legelterjedtebb –, és végül a direkt hajtás – ez közvetlen a motor forgórészére rögzített viszonylag

könnyű lemeztányért jelent. A motor lehet váltakozó áramú, mely többnyire a szabványos hálózatban alkalmazott frekvenciára épít a forgás tekintetében, vagyis szinkron motor, és lehet úgynevezett DC (egyenáramú) motor is. A közvetlen hajtású lemezjátszók esetében a legpontosabb fordulatszám elérése végett kvarcvezérlést alkalmaznak. Egyes típusoknál egy stroboszkóp lámpa segíti a fordulatszám beállítását, egy potenciométer beiktatásával, melyet általában az egyenáramú motorral szerelt típusok esetében alkalmaznak. A három megoldás közül a kvarcvezérlésű direkt hajtás a legpontosabb megoldás.

A hangkart mechanikusan áthelyezni képes lemezjátszók a lemeztányér alatt elhelyezett mechanikus szerkezet segítségével teszik mindezt. Létezik belőlük félautomata és automata kivitel is. A hangkarban helyet foglaló (sok esetben leszerelhető, cserélhető) shell biztosítja a lemeztű számára kialakított hangszedő fej rögzítését. A szögkiegyenlítés végett – nem gyakran ugyan – alkalmaznak úgynevezett tangenciális hangkart, melynél nincs ilyen jellegű probléma. A nem tangenciális esetben a fej valamekkora szöget zár be a hangkarral, melyet vagy a hangkar meghajlításával, vagy a shell kialakításával érnek el.

Hangszedőből háromféle alaptípus létezik, melyek működésükben különböznek egymástól.

Az egyik – régebben viszonylag olcsó és hatékony megoldást biztosító – típus a kristály hangszedő. Ez működését tekintve a gyémánt vagy zafír tű által a lemeztől letapogatott rezgéseket egy vékony cső segítségével egy piezokristályra vezeti, mely a rezgést elektromos jellé alakítja. Ez a letapogató rendszer viszonylag nagy túnyomással működik jól, és ez sajnos nem kíméli a mikrobarázdás lemezünket. Hangzás tekintetében nem túl jó, de kielégítő eredmény elérhető vele. Nagy előnye, hogy nem igényel külön korrekciós előerősítő fokozatot, viszont impedanciaillesztés szempontjából nagyimpedanciás bemenetet igen.

A másik szinte a legelterjedtebb, már jó hangzást biztosító megoldás a mozgó mágneses hangszedő. Ez esetben (sztereó kialakítást feltételezve) két egymással 90°-os szöget bezáró tekerccs között mozgatja a vékony cső végére rögzített mágnest a lemeztől letapogatott mechanikai rezgés. Túnyomás tekintetében változatos megoldásokkal találkozhatunk, mert van itt  $\frac{3}{4}$  grammtól kezdődően kristály hangszedőt megszügyenítő 5 grammos nyomást igénylő kialakítás is. Itt már kell egy – a köznapi életben alkalmazott erősítők nagytöbbségében az alkalmazás idő intervaluma alatt fellelhető – korrekciós előerősítő fokozat, mely a mély hangokat kiemeli, a közép hangokkal nem tesz semmit, és a magas hangokat pedig csillapítja. Erre

azért van szükség, mert a lemez előállításakor így rögzítik a jelet – a barázda szélességének csökkentése miatt (több hanganyag fér el egy lemezen). Viszonylag egyszerűen cserélhető az elkopott tű, melyet még manapság is forgalmaznak bizonyos helyeken.

A harmadik megoldás a mozgó tekercses megoldás. Ez csak fejjel együtt cserélhető, nagyon drága, és hangzásvilágát tekintve ez a legjobb minőségű. Az előerősítő fokozat ide is kell, de az erősítésének nagyobbak kell lennie az előzónél, mert a jel feszültsége csupán kb. az egytizede a mozgó mágneses hangszedőének. Létezik olyan erősítő, amely mindkét jelet képes fogadni. A megoldások közül a legjobb hangminőség a mozgó tekercses hangszedővel érhető el, de ára, és a szükséges további eszközök miatt inkább a mozgó mágneses hangszedőt használjuk.

#### 4.2.1.2. Várható zajok

A hanglemezek mechanikus adathordozók, ezért nagyon sérülékenyek. A nem megfelelő tárolás következtében a por komoly károkat tud okozni a hangzásban. A nem megfelelő poreltávolítás apró karcokat okoz, melyek hallhatók lesznek lejátszáskor. A lemez felületére az ujjainkról került zsír is károsítja a hanghordozót, továbbá a felületen nagyobb karcok is kialakulhatnak (a nem megfelelően mozgatott hangkar következtében a tű is okozhat ilyen karcokat), melyek az impulzív zajokért felelősek.

A hanglemez kialakítása során a felületi egyenetlenségek keletkeznek, melyek szintén hallhatóvá válhatnak lejátszáskor. Ezért van különbség a keleti és a nyugati nyomású lemezek között, mivel más a gyártás alapanyaga és némileg a technológiája is.

A lemezejátszó mechanikája is képes zajokat előállítani. A lemeztányér a csapágyazás és kenés függvényében valamilyen mértékben rezeg. Ez egy precízen csapágyazott motor esetén egy direkt hajtású lemezejátszóban nem okoz problémát, mert általában a hallható tartományon kívül esik a rezgés frekvenciája.

Különbséget kell tenni a lemez száraz, illetve nedves lejátszása között is. A nedves lejátszás esetén a lemezt mosófolyadékkal lejátszás közben mossuk, így a szennyeződések eltávolítása is megtörténik, továbbá a folyadék vékony réteget képez a tű és a barázda információt tartalmazó részei között, ezért a hangzás is változhat nagyon kis mértékben.

A lemez kivitelezésénél fogva a korrekciós előerősítő fokozat is zajt termel, majd az azt követő illesztő fokozat viselkedik zajgenerátorként. Ezekben az erősí-

tőkben az aktív elemek melegszenek, termikus zajt előállítva. Továbbá a nagy erősítés és a mágneses működés miatt érzékenyek az elektromágneses zajokra is, mely megfelelő szigeteléssel orvosolható.

Tehát a lemezen megjelenik a helyreállítandó hézagokat maga után vonó impulzív zaj, és a spektrumát tekintve nemfehér zaj is, melyeket két különböző eljárással tudunk csökkenteni.

#### 4.2.1.3. Zajcsökkentési eljárás

A felvétel elkészítésekor figyelni kell az alkalmazott technikák jelszint és impedancia-illesztésére, valamint a mintavételezési frekvenciára és a kvantálási hossza is. Lehetőleg a cél adathordozóéval kompatibilis értékeket kell választani, a minőségromlás elkerülése érdekében.

A lemezjátszók készülhettek a keleti, illetve nyugati régióban is. A keleti régió általában az ötpólusú tuchel aljzatokat alkalmazta a jelek továbbítására, amely nem csak kialakításában, hanem jelszintjében és impedanciájában is eltér a nyugati régióban elterjedt (4 darab) RCA csatlakozókhöz képest. A tuchel csatlakozók kisebb jelszinten dolgoznak, és nagyobb bemeneti impedanciát igényelnek, továbbá kisebb felületen érintkeznek, tehát nem alkalmasak közvetlenül digitalizálásra. (Természetesen egy egyszerű erősítőfokozat közbeiktatásával alkalmassá tehetők a feladat elvégzésére az ilyen berendezések is.) A célszerű választás tehát a folyamat kivitelezésére egy *nyugati gyártású lemezjátszó, mozgó mágneses* hangszedővel, a megfelelő előerősítő fokozatot tartalmazó szintén *nyugati gyártású erősítővel*, mely illeszthető a szabványos hangkártyák vonalbemenetére.

A folyamat a digitalizálással kezdődik. A feladatra a számítógépbe célszerű egy eleve jó értékű jel–zaj aránnyal rendelkező, kis torzítású hangkártyát választani. A Creative cég régebben is készített ilyeneket, és ma is gyárt alkalmas eszközöket. (A Sound Blaster sorozat elemei.) A digitalizáláskor a jel erős zajszenyezéssel érkezik az eszközbe.

Az első lépés az impulzív zajok megkeresése és eltávolítása. Erre alkalmas az előző részben említett Adobe Audition program. Itt a határértékek feltérképezése után indulhat a folyamat, sajnos az eredmény csak a végrehajtás után lesz hallgatható, nincs „előzetes bemutatás” lehetőség. A folyamat az impulzív zaj mennyiségétől és a számítógép sebességétől függően akár több órát is igénybe vehet. Közben látható, hogy hány mintát vélt zajnak, és hány mintát sikerült kijavítania a programnak. Természetesen az javított/kihagyott arány a beállítás függvénye, na-



gyon karcos lemez esetén sok impulzív zaj van, melyek javítása eredményezheti a hangminőség romlását is, oly formában, hogy a magas frekvenciájú jeleket is impulzív zajnak tekintik.

A következő lépés a spektrális kivonás. Ekkor mintát kell venni egy csak zajos részből, mely lehet a lemez bevezető része. A mintában szereplő pillanatképek száma beállítható, de nem feltétlenül eredményez jobb minőséget a nagyobb szám. Fontos paraméter a zaj csökkentésének mértéke, a precizitási szorzó és a simítás is. Ezek együttesen határozzák meg az eredmény minőségét. Van „előzetes bemutatás”, tehát meghallgathatjuk, hogy milyen eredményre számíthatunk, sőt megfigyelhető csak a zaj is.

A végén célszerű a jelet normalizálni, a hangerő (vagy jel–zaj arány) javítása érdekében. A mentés történhet szakaszonként is, ezzel a hanganyag darabolása is megtörténhet. Az eljárással akár 60–80 dB jel–zaj arány is elérhető, ami a kiindulási 30–40 dB-hez képest nem rossz eredmény.

#### 4.2.2. Hangkazetta archiválása

##### 4.2.2.1. A magnetofon és a hangszalag

A magnetofon vékony mágnesezhető réteggel bevont műanyag szalagon rögzíti a jeleket. A jelek, mivel az átmagnesezhető réteg vékony, nagyon kis feszültséget indukálnak a letapogató rendszerben, vagyis a magnetofon fejben, mely egy speciális kemény, kopásálló bevonattal rendelkező tekercs(pár). A kicsiny indukált feszültség erősítése egy korrekciós előerősítő fokozattal kezdődik, melynek a magas hangokat kell kiemelnie. A lejátszó rendszer ezt a szalagot húzza egyenletes sebességgel a fej előtt. A szalagsebességről egy kisebb átmérőjű fém és egy rugóval rászorított nagyobb átmérőjű gumigörgő gondoskodik, egyes esetekben ezt a rendszert megduplázzák, a stabilabb szalagsebesség érdekében. A görgők viszonylag könnyen szennyeződnek, ami a szalag nem kívánt oldalirányú mozgásához vezet; ez károsan befolyásolja a lejátszást. A szalagsebesség egy fontos jellemzője a berendezésnek, ami 4,76 cm/s egy átlagos magnetofon esetében. Ennek beállítása referenciaszalaggal és oszcilloszkóppal történik. A gyártók egy túréson belülre állítják ezt az értéket, tehát két azonos szalagsebességű berendezést nehéz találni. A lejátszó és felvevő kombinált fej szennyeződhet, átmagneseződhet, ami szintén minőségromlást eredményezhet. Fontos paraméter még a szalagon található sáv és a fejben található érzékelő egymáshoz viszonyított elhelyezése is. Nem megfelelő beállítással a magasabb frekvenciájú hangok mennyiségének nagymértékű csökkenése

érhető el. A gyártók ezt is határértékek közé állítják, de létezik gyárilag rosszul beállított magnetofon is.

A szalag esetén problémák a szalagra felvitt mágnesezhető réteg minőségével kezdődnek, ebből alapjában véve négyféle típust alkalmaztak. Az első a normál vasoxid, a második a króm dioxid, a harmadik a ferrokróm (ezt már nem alkalmazzák) a negyedik pedig a magnetit. A következő probléma a mágnesezhető réteg szemcsézete, egészen pontosan a mágneses információhordozók elrendezése. Ez a szalag gyártójától függően – azonos típusú mágnesezhető anyag esetén – anyagminőségben is eltérhet egymástól, tehát vannak jobb minőségű és rosszabb minőségű hangkazetták, annak függvényében, hogy a mágnesességet mennyi ideig képesek megőrizni az egyes szalagok. A harmadik probléma a műanyag réteg vastagsága, mely lejátszáskor is okozhat problémát a szalagsebesség és az oldalirányú mozgás miatt, de tárolás esetén kérdéses, hogy az egyes átmágnesezett rétegek hogyan hatnak egymásra az idő múlásával. A műanyag hosszirányban képes a megnyúlásra, ami tovább fokozza a problémákat. Továbbá a vékony szalag nagyon érzékeny a mechanikai behatásokra, mint például a gyűrődés, amely durva minőségromláshoz vezethet. Az ilyen jellegű hiányok sem pótolhatók a jel folytonosságában, még digitális módszerrel sem. Nagyon fontos szerepe van még szalagok esetén az előmágnesezésnek, ugyanis ez is potenciális zajforrás.

#### 4.2.2.2. Várható zajok

A szalagos egységek esetében az előmágnesezési zaj után a szalagnak is van egy meghatározott értékű zaja, melyet egy olyan elektronika követ, ahol több termikus zajgenerátor is működik egy időben. A zaj ezért összetett, és általában egy elég jól hallható sistergés formájában jelentkezik, tehát főleg a magasabb frekvenciákon ölt jelentős mértéket.

A magnetofon esetén alkalmazható a Dolby zajcsökkentő rendszer, de lejátszáskor csak akkor, ha a felvétel is azzal készült, ellenkező esetben nem csak a zaj, hanem a magas hangok is eltűnnek. Ha nem tudjuk egy szalag esetén, hogy használtak-e valamilyen analóg zajcsökkentőt a felvételkor, akkor a berendezés zajcsökkentőjét kapcsoljuk ki, mert így a digitális zajszűrő hatékonyabban távolíthatja el a zajt, és több hasznos jel maradhat épen.

A magnetofonnal – a gyártók ígéretei szerint – akár 60 dB jel–zaj arány is elérhető a Dolby zajszűrővel, de sajnos a szalagok minősége és a közbeiktatott erősítőfokozatok miatt a digitalizálás után ez az érték 40–45 dB-re romlik.

Mivel vannak olyan rögzítő eszközök ebben a kategóriában, amelyek a törlést és egyben az előmágnesezést egy állandó mágnessel valósítják meg, ezért az előmágnesezés szintje közel sem lesz állandó, sokkal inkább jól hallható alacsony frekvenciájú előmágnesezési zaj lesz belőle. Ez a zajforma oly mértékben elnyújtott, és véletlenszerű, hogy algoritmikus csökkentése esetenként nem is lehetséges, ugyanis a zaj mintavételezése során nem választható ki úgy zajos rész, hogy e zajból kellő mennyiségű és minőségű jusson a mintába megfelelő kivonáshoz.

#### 4.2.2.3. Zajcsökkentési eljárás

A magnetofonszalagokról érkező jelek főként magas frekvenciájú zajt tartalmaznak, mely a korrekciós előerősítő fokozatból adódik. Ez a zaj mintavételezést követően spektrális kivonással csökkenthető a zavaró mérték alá.

Csak olyan esetekben ajánlott a magnetofonszalagok digitális hanghordozóra másolása, ha az anyag bír akkora fontossággal, hogy megtegyük (más hanghordozón, nem hozzáférhető). A magnetofonszalagokról tudjuk, hogy a normál szalag sávátvitele alig több mint a fele a krómdioxidos vagy magnetit szalagokéhoz képest, és a szalag zaja, ami a mágneses információhordozó minőségéből adódik, is lényegesen nagyobb az ilyen szalagoknál.

A magnetofonnál általában nem jelentkezik impulzív jellegű zaj, ezért a zajcsökkentés elvégezhető egyetlen lépésben. A digitalizált jelből (az elejéről vagy a végéről) zajmintát veszünk, majd pedig spektrális kivonással távolítjuk el a zajt. Az eredmény – bár az eljárás egy lépésből áll csupán – nem lesz olyan minőségű, mint a hanglemezek esetében, ugyanis a magasabb frekvenciájú hangokat a rögzítő rendszer és a szalag már korábban eltüntette, ezért lejátszáskor sem voltak már jelen. Ezek a hangok már semmilyen eszközzel nem állíthatók vissza.

## 5. Összegzés

Dolgozatom célja volt a környezetünkben előforduló zajok osztályozása, és azok csökkentésére megoldást találni. A potenciális zajforrások nagytöbbségére a dolgozatom első szakaszában világítottam rá. Tekintettel arra, hogy zaj hasonlóan az információt hordozó jelhez véletlenszerű, ezért modellezése és csökkentése közel sem olyan egyszerű, mint ahogyan azt a köznapi ember gondolja. A zajok modellezése a zajcsökkentés szempontjából jelentős, ugyanis egy megfelelő modell segítségével az eljárás könnyebben algoritmizálható, és alkalmazható a digitális jelfeldolgozásban. A számítógépek miniatürizálásával és gyors fejlődésével korunkban egyre nagyobb teret kap a kommunikációs és jelrögzítő rendszerekben a digitális jelek alkalmazása, mely a feldolgozhatóság és a hozzáadódó zajokkal szembeni jó „ellenálló-képességének” köszönheti sikerét.

A zajok modellezését kétféle szempont szerint közelítettem meg dolgozatomban, melyekből kiderült, hogy a véletlenszerű jelek előre meg nem jósolható értékeket is felvehetnek. A modellek alkalmazhatók más köznapi véletlenszerű események modellezésére is, melyek messze esnek az informatika tudományterületétől, mint például a közgazdaságtan, vagy a meteorológia, vagy az orvostudomány.

A probléma megoldása a zajcsökkentési eljárások alkalmazása. Többféle módszer létezik, melyekből csak néhányat volt alkalmam kiemelni, dolgozatom szűkös keretei miatt. A módszereket az adott problémának megfelelően, jelen esetben a zajok fajtáihoz illően kell alkalmazni, a várt eredmény elérése érdekében. A zajok csoportosítása, és forrásuk meghatározása ezért nagyon fontos, a legalkalmasabb eljárás kiválasztása szempontjából. Egy véletlenszerű jelsorozatból bizonyos esetben nehéz eldönteni még a szétválogatás után is, hogy melyik a számunkra hasznos információt hordozó jel. Erre nagyon szemléletes példa egy gépjármű motorja, mely az utas szempontjából zajforrás, de egy hozzáértő szerelő sok értékes információhoz juthat a hallott hangokból.

Az utolsó szakaszban az elméleti ismeretek gyakorlatba ültetését próbáltam kifejteni, saját tapasztalataim alapján. Több mint tíz évvel ezelőtt volt alkalmam több „házi stúdióban” készült könnyűzenei demófelvételhez segítséget nyújtani, és jómagam is aktív részese voltam egyes felvételek elkészítésének, mint hangszeres „zenész”. Jelenlegi iskolámban néha alkalmam nyílik kamatoztatni korábbi ismereteket.

reteimet, és dolgozatom megírásával nagyobb rálátást nyerhettem a későbbi munkámhoz a zajok forrásaira és megfelelő mértékű csökkentésükre.

A régebbi hanghordozókon tárolt, többnyire zenei felvételek korszerű adathordozóra mentése több ismerősömnek eszébe jutott már, de megfelelő eszköz hiányában maguk nem tudták elvégezni a műveletet. Érdeklődési körömbé tartozik a hangrendszerek felépítésének, és működésének megismerése és tanulmányozása, ezért több, mikrobarázdás lemez lejátszására alkalmas eszköz és magnetofon is található otthonunkban, mellyel a hanganyagok digitalizálásához forrást tudok biztosítani. A számítógépben található hangkártyákra mindig figyeltem, hogy ne csak gyenge minőségű hangjelet tudjon produkálni a komolyabb hangrendszer számára, hanem megfelelő működésével minőségi zenehallgatásra legyen alkalmas.

A dolgozat megírása során törekedtem arra, hogy a címben megragadott témát minél jobban, és több szempont szerint is kifejthessem, amit végeredményében sikerült megvalósítanom.

## 6. Irodalomjegyzék

1. Saaed V. Vaseghi – Advanced Digital Signal Processing and Noise Reduction  
Second Edition  
Copyright © 2000 John Wiley & Sons Ltd.
2. Valószínűségszámítás és statisztika előadások anyaga  
Nagy Márta – 2001–2002 tanév
3. Molnár-Sáska Gábor – Rejtett Markov modellek statisztikai vizsgálata  
Tézisfüzet – 2005  
forrás: [http://doktori.math.bme.hu/Molnar-Saska\\_tezisfuzet.pdf](http://doktori.math.bme.hu/Molnar-Saska_tezisfuzet.pdf)
4. dr. Tolnai János – Rádióamatőr vizsgára felkészítő tananyag  
Puskás Tivadar Távközlési Technikum, 2006.  
forrás: [http://www.puskas.hu/r\\_tanfolyam/r\\_tananyag.html](http://www.puskas.hu/r_tanfolyam/r_tananyag.html)

Debreceni Egyetem  
Informatika Kar

# Szakedolgozat

## Zajok és eltávolításuk

**Témavezető:** Dr. Papp Zoltán  
egyetemi adjunktus

**Készítette:** Gambár Péter István  
III. informatika tanár szakos,  
levelező tagozatos hallgató

Debrecen  
2009.

# Tartalomjegyzék

1. Bevezetés.....	4
1.1. A cím magyarázata.....	4
1.2. Témaválasztás indoklása .....	4
2. Zajok és modellezésük .....	5
2.1. Zajok fajtái és forrásai.....	5
2.1.1. Általános felosztás .....	5
2.1.2. Fehér zaj.....	7
2.1.3. Színezett zaj .....	7
2.1.4. Impulzív zaj.....	8
2.1.5. Tranziens zaj impulzus.....	10
2.1.6. Termikus zaj .....	11
2.1.7. Schottky-zaj.....	12
2.1.8. Elektromágneses zaj .....	13
2.1.9. Csatornatorzítások.....	14
2.1.10. Zajmodellezés .....	15
2.1.10.1. Hozzáadódó fehér normális eloszlású zajmodell (AWGN).....	16
2.1.10.2. Rejtett Markov modell.....	16
2.2. Zajmodellek .....	18
2.2.1. Valószínűségi modellek.....	18
2.2.1.1. Véletlenszerű jelek és sztochasztikus folyamatok.....	18
2.2.1.2. Valószínűségi modellek .....	19
2.2.1.3. Állandó és nem állandó véletlen folyamatok .....	23
2.2.1.4. Véletlen folyamatok várható értéke.....	24
2.2.2. A rejtett Markov modell.....	29
2.2.2.1. Nem stacionárius folyamatok statisztikai modellje .....	29
2.2.2.2. Rejtett Markov modellek.....	31
2.2.2.3. A rejtett Markov modell betanítása .....	34
3. Zajsűrítési módszerek.....	36
3.1. Zajzár .....	36
3.2. Analóg valós idejű zajsűrítők.....	36
3.2.1. Dolby zajsűrítő.....	36
3.2.2. Zajérvénytelenítés.....	38
3.3. Digitális jelek feldolgozása.....	39



---

3.3.1. Átalakítások.....	39
3.3.1.1. Analóg–digitális átalakítás.....	39
3.3.1.2. Digitális–analóg átalakítás.....	40
3.3.2. A Wiener szűrő.....	42
3.3.2.1. A legkisebb négyzetek becslése.....	42
3.3.3. Zajcsökkentés spektrális kivonással.....	44
3.3.3.1. A spektrális kivonás módszere.....	45
3.3.3.2. A spektrális kivonó megvalósítása.....	47
3.3.4. Az impulzív zaj eltávolítása.....	50
3.3.4.1. Impulzív zaj matematikai modellje.....	50
3.3.4.2. A medián szűrő.....	51
3.3.4.3. Impulzív zaj eltávolítása lineáris prediktív modell segítségével ...	53
4. Zajszűrés a gyakorlatban.....	58
4.1. Zenekari stúdiómunka.....	58
4.1.1. A felvétel.....	58
4.1.2. Az utómunka.....	59
4.2. Hagyományos hanglemez, hangkazetta archiválása CD-re.....	61
4.2.1. Hanglemez archiválása.....	61
4.2.1.1. A lejátszó rendszer.....	61
4.2.1.2. Várható zajok.....	63
4.2.1.3. Zajcsökkentési eljárás.....	64
4.2.2. Hangkazetta archiválása.....	65
4.2.2.1. A magnetofon és a hangszalag.....	65
4.2.2.2. Várható zajok.....	66
4.2.2.3. Zajcsökkentési eljárás.....	67
5. Összegzés.....	68
6. Irodalomjegyzék.....	70

## 1. Bevezetés

### *1.1. A cím magyarázata*

Általában hangfelvételek (függetlenül attól, hogy milyen eszközzel készülnek és milyen adathordozóra) és kommunikációs rendszerek esetében felütik fejüket az additív zajok. Ezen zajokat a hasznos jelhez szokták viszonyítani, és sajnos adott esetben a hasznos jel nagyságát is meghaladhatják, mely nagyban károsítja a jelek értelmezhetőségét. Több esetben azonban lehetőség nyílik az ilyen zajok mértékének, erősségének csökkentésére, mind hagyományos, mind digitális rendszerben. A zajok eltávolításánál nagyon fontos, hogy a hasznos jel minél nagyobb hányada maradjon ép, értelmezhető, és lehetőleg torzításmentes. Zajforrástól függően a zajok csökkentése hagyományos úgynevezett analóg rendszerekben is megvalósítható, akár megfelelő alkatrészek és eszközök alkalmazásával is. Jelen dolgozatban a tárolt és szállított hangjelek zajainak felismerésével, fajtaival és azok csökkentésével foglalkozom.

### *1.2. Témaválasztás indoklása*

Több éve egy zenei általános iskolában tanítok, ahol esetenként szükség van a kórus munkájának megörökítésére nem csak képi, de hang formájában is. A hangfelvételek megfelelő minőségű elkészítése és tárolása érdekében a rögzítő technika és az előtte található átalakító rendszerek által termelt zajok eltüntetése adott esetben nagyon fontos. Célom, hogy a dolgozat megírásával iskolámban minél hatékonyabb munkát tudjak végezni ebben a tekintetben is.

Személyes érdeklődési körömbé tartozik a zenei hangfelvételek digitális feldolgozása, illetve az erre alkalmas berendezések és szoftverek megismerése is, tehát többek között ez is motivált a szakdolgozat témájának megválasztásánál.

## 2. Zajok és modellezésük

### 2.1. Zajok fajtái és forrásai

#### 2.1.1. Általános felosztás

A zaj definiálható olyan nem kívánt jelnek, mely megakadályozza (zavarja) egy másik jellel (hasznos jel) történő kommunikációt vagy a másik jel mérését. A zaj maga is egy információhordozó, a zajforrás szempontjából. Például egy autó motorjának zaja információval szolgál a motor állapotáról. A zajok igen sokfélék és változatosak lehetnek, kezdve a hangfrekvenciás akusztikus zajoktól a rádiófrekvenciás elektromágneses zajokig.

A másik említésre méltó probléma a hasznos jel kiindulási alakjától való nem kívánatos eltérése, torzulása. Ebbe a csoportba tartoznak a különféle visszhangok, felharmonikusok, és az elveszett jelerészetek.

A zajok és torzulások a kommunikációs és mérési rendszerek fő ellenségei, ezért e káros mellékhatások eltüntetése érdemel kiemelt figyelmet. A zajszűrést és torzítás kiküszöbölését főként a mobil kommunikációban, a beszéd felismerésben, orvosi jelfeldolgozásban, radar és hanglokátor rendszerekben alkalmazzák a köznap életben.

A zajok életünk minden területén, minden környezetben jelen vannak, különféle formában. Az zaj forrásától függően a következő nagy csoportok különíthetők el egymástól:

- a.) *akusztikus zaj*: leggyakrabban mozgásból, vibrációból (rezgésből), ütközésből származik, ezáltal ez a zajforma – eltérő mértékben ugyan – jelen van a mindennapi környezetünkben. Ezt a zajt legtöbbször haladó gépjárművek, légkondicionálók, számítógépek hűtőventillátorai, közlekedés, emberi beszélgetések, szél, eső, stb. okozzák.
- b.) *elektromágneses zaj*: minden frekvencián jelen van, különösen a rádiófrekvenciákon. Minden elektromos berendezés, például a rádió- és TV adóállomások, és vevőkészülékek állítanak elő jelentős elektromágneses zajt.
- c.) *elektrosztatikus zaj*: feszültség jelenléte okozza, áram nélkül vagy folyó árammal. Ennek a zajnak az egyik legjelentősebb forrása a fénycsővilágítás.

d.) *csatornatorzítás, visszhang, és csillapítás*: az átviteli csatorna nem ideális karakterisztikája következtében történik. A rádiófrekvenciás csatornák, melyeket a mobil kommunikációban használnak, – mikrohullámú frekvenciákon – különösen érzékenyek a terjedés karakterisztikájára és környezetére.

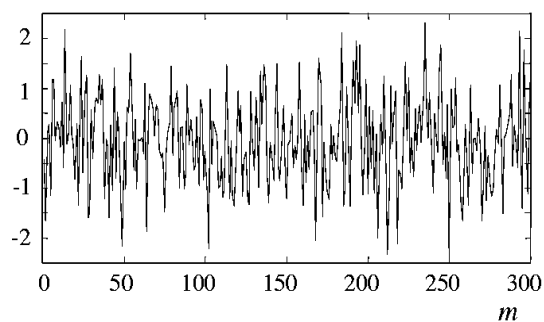
e.) *feldolgozási zaj*: az a zaj, amelyet a jel digitális–analóg átalakítása eredményez, például a zaj digitálisan kódolt beszédben vagy hasonló jelekben, vagy az elveszett csomagok a digitális adatkommunikációs rendszerekben.

A zaj frekvenciájának vagy időkarakterisztikájának függvényében a következő kategóriákba csoportosítható:

- *Keskeny hullámsávú zaj*: olyan keskeny hullámsávon érkező zajhatás, mint az 50 Hz-es vagy 60 Hz-es brummogás (búgás) mely többnyire az elektromos tápegységekből, váltakozó áramú vezetékekből származik.
- *Fehér zaj*: tisztán véletlen zaj, melynek teljesen egyenes az energiaspektruma. A fehér zaj elméletileg az összes frekvenciát tartalmazza, azonos intenzitással.
- *Sávkorlátozott fehér zaj*: egyenes spektrumú zaj korlátozott sáv szélességgel, mely általában az alkalmazott eszköz korlátozott szélességű sávját, vagy az általunk használt sávot jelenti.
- *Szinezett zaj*: nemfehér zaj, vagy egyéb szélessávú zaj, melynek nem egyenes a spektruma, például: rózsaszín zaj, barna zaj, autoregresszív zaj.
- *Impulzusszerű zaj*: véletlenszerű rövid időtartamú impulzusokból áll, véletlenszerű amplitúdóval, hosszal és felbukkanással.
- *Tranziens zaj*: viszonylag hosszú időtartamú impulzusokból áll.

### 2.1.2. Fehér zaj

A fehér zaj definíció szerint egy összefüggéstelen jelekből álló zaj, mely minden frekvencián azonos energiájú. A zaj, melynek energiája minden frekvencián azonos  $\pm \infty$  tartományban, szükségképpen végtelen nagy energiájúnak kellene lennie, ennek következtében ez csak egy elméleti fogalom. Habár egy sávkorlátozott zaj,



1. ábra Fehér zaj

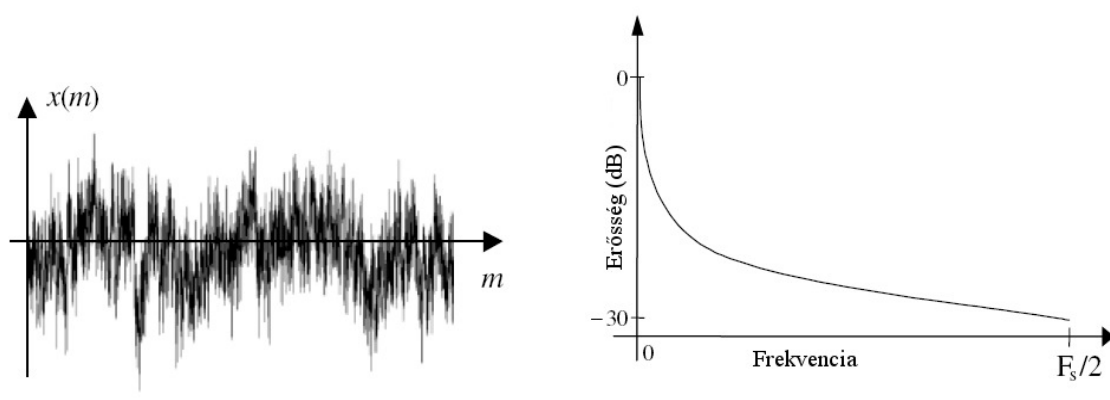
egyenes spektrummal – ha a zaj a teljes vizsgált sávot lefedi, amit a kommunikációs rendszer biztosít – a rendszer szempontjából minden tekintetben fehér zajnak számít. Például egy 10 kHz sávátvítellel rendelkező hangrendszer számára minden 10 kHz-nél szélesebb sávú egyenes spektrumú hallható zaj fehér zajnak

tűnik.

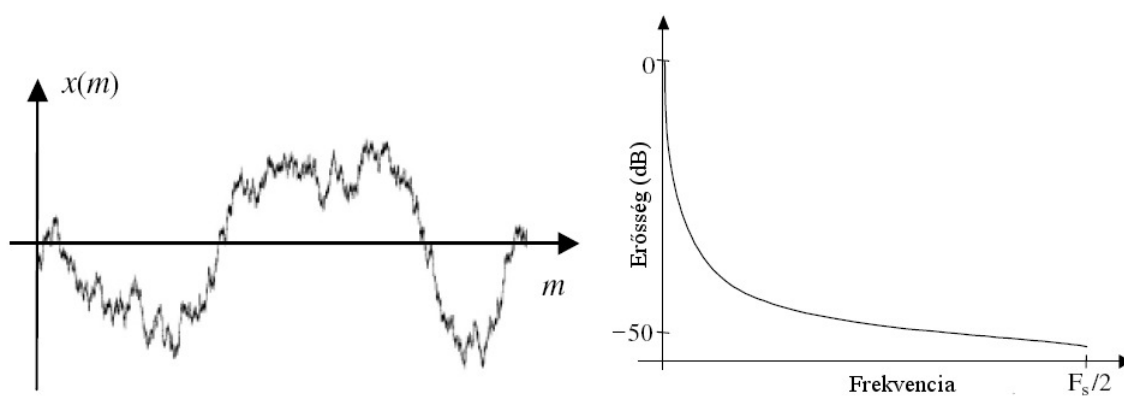
A tiszta fehér zaj egy elméleti fogalom, mivel végtelen energiájúnak kellene lennie a végtelen frekvenciatartomány lefedésére. Ráadásul a diszkrét időpontokban mintavételezett jelnek szükségképpen sávkorlátozottnak kell lennie, ahol a legnagyobb frekvenciaérték kisebb, mint a mintavételezési frekvencia fele. A sávkorlátozott fehér zaj egy sokkal gyakorlatibb fogalom, mely definíció szerint egy egyenes spektrumú zaj egy adott frekvenciasávban. Egy ilyen zaj energiája nem végtelen, ezért kiszámítható.

### 2.1.3. Színezett zaj

Bár a fehér zaj elmélete meglehetősen valós, és matematikailag kényelmes és használható közelítést ad több zajjal kapcsolatban, melyek felütik fejüket a telekommunikációs rendszerekben, sok egyéb zaj viszont nem fehér. A színezett zaj kifejezés bármely olyan szélessávú zajra utal, melynek spektruma nem a fehér zajéval egyező. Például a legtöbb hangfrekvenciás zaj, mint a mozgó autók zaja, a számítógépek ventilátorainak zaja, a fűrés zaja és emberi beszélgetés zaja, nemfehér, és túlnyomórészt alacsonyfrekvenciájú hangokat tartalmaz. Egy fehér zaj adott csatornán történő továbbításakor a zaj bizonyos frekvenciákon történő csillapodásával jellemezhető a „színezett zaj”, a csatorna spektrumának alakjától függően. A színezett zaj két klasszikus változata a rózsaszín (2. ábra) és a barna (3. ábra) zaj.



2. ábra Rózsaszín zaj és spektruma



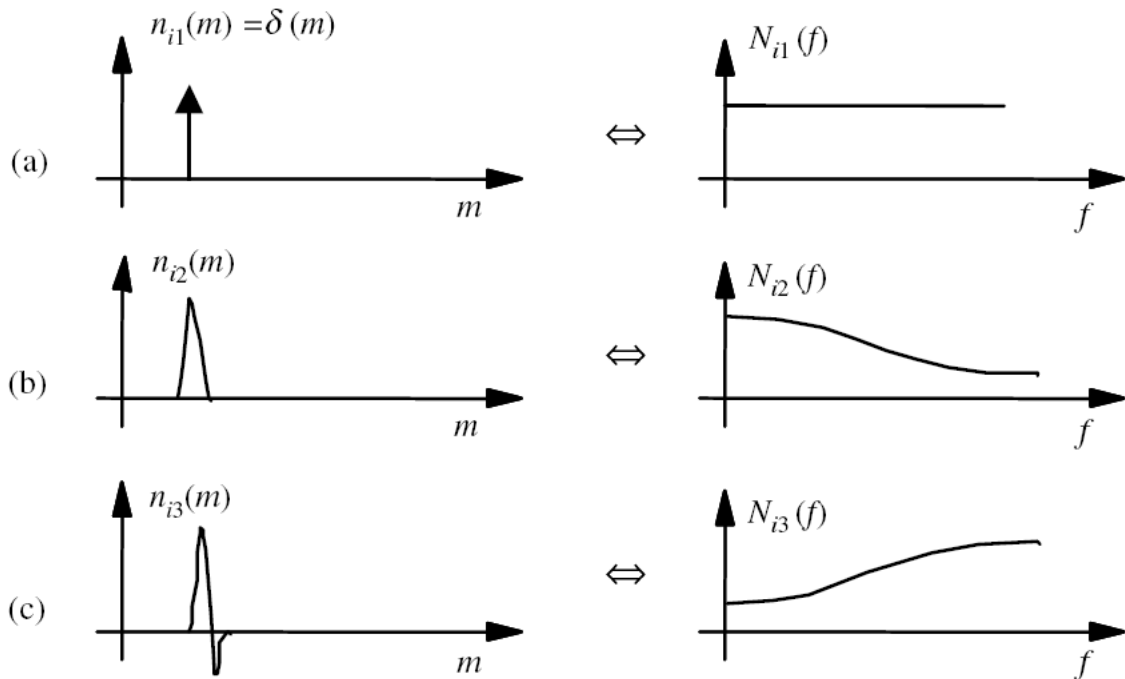
3. ábra Barna zaj és spektruma

#### 2.1.4. Impulzív zaj

Az impulzív zaj rövid időtartamú „be/ki” impulzusokból áll. Ezek a zajok többféle forrásból származhatnak, mint például kapcsolási zaj, kedvezőtlen csatorna-környezet, kimaradás vagy felületi kopás a hanghordozón, számítógép-billentyűzet kattogása, és így tovább. A 4. ábra (a) része egy ideális impulzív zajt mutat, annak spektrumával. A kommunikációs rendszerekben előforduló impulzív zaj jellemző paramétere a hossza, mely normális esetben több mint egy minta hosszúságú. Például hangrendszerek esetében a rövid időtartamú, éles impulzusok lehetnek akár 3 milliszekundum (60 minta 20 kHz-es mintavételezési frekvencia esetében) hosszúak, melyeket impulzív zajnak tekinthetünk. A 4. ábra (b) és (c) része két példát mutat rövid időtartamú impulzusokra és azok spektrumára.

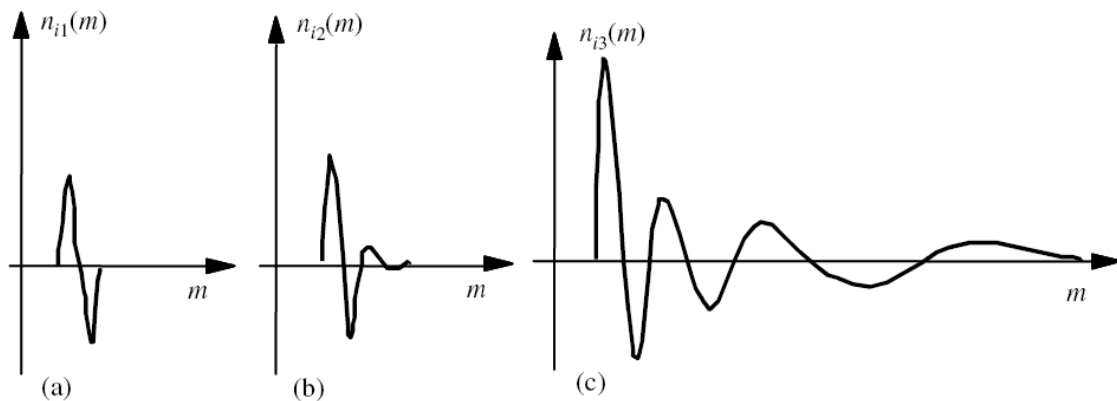
Egy kommunikációs rendszerben az impulzív zaj a térben és időben több pontból ered, azután a csatornán keresztül halad a vevő felé. A vett zaj időben szétszórta és a csatorna által alakított, és úgy kezelhető, mint a csatorna impulzus jelre adott reakciója. Általában a csatornák karakterisztikája lehet lineáris vagy nem lineáris,

időben állandó, vagy időben változó. Azonkívül sok kommunikációs rendszer válaszul a nagy amplitúdójú impulzusokra nem lineáris karakterisztikát mutat.



4. ábra Impulzív zaj

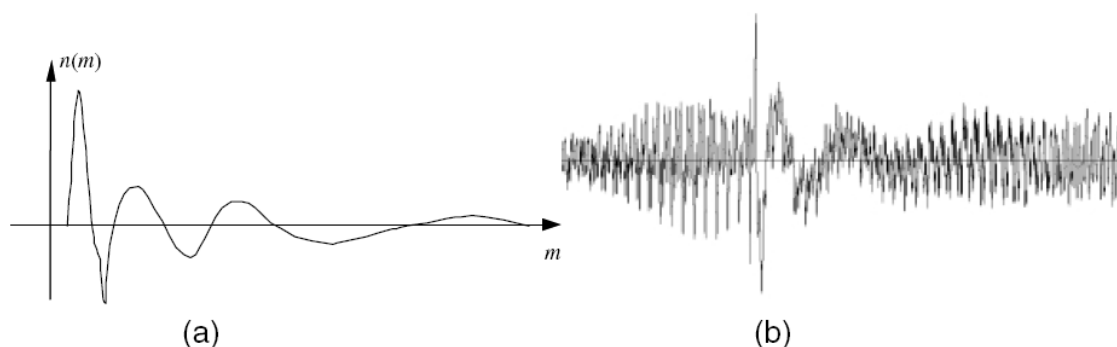
Az 5. ábra néhány példát mutat impulzív zajra, melyek tipikusan régi hanglemezeken fordulnak elő. Ebben az esetben a kommunikációs csatornának a lejátszó rendszert tekintjük, melyet időben állandónak feltételezhetünk. Az ábra három változatát mutatja a zaj-amplitúdóváltozásának az idő függvényében. Az 5. ábra (c) eleme mutatja, hogy egy nagy impulzussorozat csillapodó tranziens impulzussorozatot hoz létre. Ezek a példák a lejátszó rendszer nem lineáris karakterisztikáját jelzik.



5. ábra Példák impulzív zajra

### 2.1.5. Tranziens zaj impulzus

Az tranziens zaj gyakran áll egy viszonylag rövid éles impulzusból, melyet egy csillapodó viszonylag alacsony frekvenciájú oszcilláció követ, ahogy azt a 6. ábrán is láthatjuk. A kezdő impulzust gyakran valamilyen belső vagy külső interferencia okozza, míg az oszcillációt gyakran a csatorna rezonanciája, melyet a kezdeti impulzus gerjeszt, és tekinthető a csatorna válaszának a kezdeti impulzusra. Egy telekommunikációs rendszerben az impulzív zaj a térben és időben több pontból származik, azután a csatornán keresztül halad a vevő felé. A vett zaj időben szétszórt és a csatorna által torzított, alakított és úgy tekintendő, mint a csatorna egy impulzusra adott válaszreakciója. Az impulzív zaj a kezdő impulzussal és az azt követő oszcillációval jellemezhető, a csatorna szempontjából.



6. ábra Egy karc impulzus és zene egy hanglemezről (b), hanglemez karc impulzusának átlagolt profilja (a)

A 6. ábra jól szemlélteti a tranziens zaj impulzus alakját, tekintsük meg a karcolás impulzust<sup>1</sup> az ábrán, mely egy sérült hanglemezről származik. A karcolás impulzusok a lemeztű és a hozzá kapcsolt elektromechanikus lejátszó rendszer akusztikus válasza a lemezen található fizikai megszakítások jelenlétére. Mivel a karcolás hatására előálló tranziens zajok tulajdonképpen a lejátszó rendszer impulzus válasza, elvárható, hogy az adott rendszer különböző helyeken található impulzusokra hasonló karakterisztikát mutasson. Ahogyan a 6. (a) ábra is mutatja egy tipikus karc impulzus hullámalak gyakran két különböző részre tagolódik:

- egy bevezető nagy amplitúdójú lökésre, mely a lejátszó rendszer válasza a rögzített médiumon (lemezen) található karcolásra, melyet
- csillapodó hullámzások követnek, további torzítást okozva ezzel.

<sup>1</sup> Karc(olás) impulzus alatt egy hanglemezen, a használat során kialakult (többnyire a sávval valamilyen szöveget bezáró) karc (tekinthető pillanatnyi jelfolytonosság-hiánynak is) hatására létrejövő lejátszáskor jelentkező ugrásszerű lökést érthetünk, mely közvetlenül a mechanikus jelérzékelőre hat, és a mechanikai–elektronikai átalakítást követően végighalad a teljes elektronikus rendszeren is.



A kezdő impulzus általában viszonylag rövid, és a hossza a 1–5 ms tartományba esik, míg a hullámzó farok hosszabb időtartama elérheti akár az 50 ms-ot vagy többet is.

A 6. (b) ábrán látható csillapodó hullám (farok) frekvenciája is csökken az idő múlásával. Ez a viselkedés az elektromechanikus lejátszó-rendszer nem lineáris válaszmechanizmusa a lemezen található karcokra. Tapasztalat alapján több hanglemezről származó recsenés zaja jól meghatározható hullámalakot mutat, és viszonylag kevés tipikus sablonnal jellemezhető.

### 2.1.6. Termikus zaj

A termikus zaj – más néven Johnson-zaj (felfedezőjéről J. B. Johnsonról<sup>1</sup> elnevezve) – hővel stimulált részecskék rendezetlen mozgásából adódik. A termikus zaj elméletének alapjai a termodinamikában keresendők. A termodinamikában a szabad részecskéket hőmérsékletfüggő rendezetlen mozgás jellemzi, mint például a szabad gázmolekulák egy tartályban, vagy a szabad elektronok egy vezetőben. Bár ezen rendezetlen mozgások átlagértéke közel nulla, az átlag körüli hullámzások alkotják a termikus zajt. Például a gázmolekulák rendezetlen mozgásai és ütközései egy behatárolt térben rendezetlen hullámzást okoznak az átlagos nyomás körül. A hőmérséklet növekedésével a molekulák vagy más elemi részecskék kinetikai energiája növekszik, és ezzel a termikus zaj is.

Ehhez hasonlóan egy elektromos vezetőben nagy számú szabad elektron van jelen (az ionokkal együtt), melyek rendezetlenül rezegnek az egyensúlyi pozíciójuk körül, és útját állják a további elektronok mozgásának. Az elektronok rendezetlen szabad mozgásai spontán áramokat, vagy más szóval termikus zajt képeznek, melynek átlaga nulla, mivel feszültség hiányában minden elektron mozgása különböző irányú (nincs egyirányú áramlás). A környezet hőmérsékletének növekedésével a vezető hőmérséklete is növekszik, mely az elektronokat magasabb energiaszintre viszi, ezzel növelve a rendezetlen áramfolyásokat. Egy fémréteg ellenállásban a pillanatnyi feszültségek négyzetének átlagértéke a termikus zaj miatt a következő összefüggéssel adható meg:  $\overline{v^2} = 4kTRB$ , ahol  $k = 1,38 \cdot 10^{-23} \frac{\text{J}}{\text{°K}}$  a Boltzmann-állandó,  $T$  az abszolút hőmérséklet °K-ben,  $R$  az ellenállás ohmban, és  $B$  a sávszélesség. Az iménti egyenletből és az azt megelőző érvelésből egy áramkörü

<sup>1</sup> John Bertrand “Bert” Johnson (1887–1970): svéd születésű amerikai villamosmérnök és fizikus. Ő fejtette ki először részletesen a kábeleken közvetített információ véletlen interferenciájának alapjait. (Wikipédia – [http://en.wikipedia.org/wiki/John\\_B.\\_Johnson](http://en.wikipedia.org/wiki/John_B._Johnson))

lapba ültetett fémréteg ellenállás úgy viselkedik, mint egy termikus zajgenerátor, mely feszültségnégyzetének átlaga  $\overline{v^2}$ , és a belső ellenállása  $R$ . Az ily módon előálló termikus zajgenerátor teljesítménye számítható ki a következő összefüggéssel:

$$P_N = \overline{i^2} R = \left( \frac{v_{rms}}{2R} \right)^2 R = \frac{\overline{v^2}}{4R} = kTB [\text{W}], \text{ ahol } v_{rms}^1 \text{ az effektív feszültség értéke. A ter-}$$

mikus zaj spektrumsűrűsége megadható a  $P_N(f) = \frac{kT}{2} \left[ \frac{\text{W}}{\text{Hz}} \right]$  összefüggéssel. Ebből

az egyenletből kiderül, hogy a termikus zaj spektrumsűrűsége egyenes, azaz a termikus zaj egy fehér zajnak tekinthető (mely általában sávkorlátozott). Ezt az összefüggést különösen a magas rádiófrekvenciák ( $10^{13}$  Hz) esetében fontos.

### 2.1.7. Schottky-zaj

A Schottky-zaj<sup>2</sup> elmélete az elektroncsövek katódjának elektronkibocsátás-ingadozásából származik. Az különállónak tekintett elektronok áramlás során véletlenszerű időpillanatokban érkeznek, ezért az átlagérték körüli ingadozást okoznak a részecskefolyamban. Ezek az ingadozások a részecskefolyam sebességében Schottky-zajt okoznak. Egyéb példák Schottky-zajra a fotonok áramlása egy lézer fénysugárban, elektronok és lyukak áramlása valamint rekombinációja<sup>3</sup> félvezetőben, vagy a fotóelektronok árama, melyek a fotódiódákban emittálódnak. A véletlenszerűség jelenléte a részecske-kibocsátás vagy -érkezés mértékében magába foglalja, hogy a Schottky-zaj Poisson eloszlással modellezhető. A hosszú megfigyelési idejű kísérletek esetén az érkezések átlagérték körüli ingadozása a normális eloszláshoz közelít. Míg a termikus zaj a részecskék „nyugalmi állapotban” történő véletlenszerű mozgásából adódik, addig a Schottky-zaj egy (feszültséggel) irányított részecske-áramlásra értelmezhető.

<sup>1</sup>  $v_{rms}$  – root mean squared voltage: a feszültség effektív értéke

<sup>2</sup> Walter Hermann Schottky német fizikus, a film-rácsú vákuumcső (1915) és a tetróda (1919) feltalálója – munkáit a Siemensnél végezte. 1938-ban megfogalmazott egy elméletet, mely a Schottky-hatást jelezte, ma pedig a Schottky-diódákban használjuk. – [http://en.wikipedia.org/wiki/Walter\\_H.\\_Schottky](http://en.wikipedia.org/wiki/Walter_H._Schottky)

<sup>3</sup> „A félvezető nem csak szabad elektronjai útján vezeti az áramot (mint a fémek), hanem egy másik vezetési mechanizmus: a lyukvezetés útján is. Az eredetileg elektromosan semleges atom a negatív töltésű elektron távozása folytán pozitív töltésűvé válik, ezért a lyukat pozitív töltésűnek tekintjük. Előfordulhat, hogy egy szomszédos atom valamelyik vegyértékelektronja szintén elszabadul, és „beugrik” az előbb említett (pozitív töltésénél fogva az elektront vonzó) lyukba. Ezzel az előbbi lyuk (és az oda beugró szabad elektron is) megszűnik, hiszen a rácsszerkezet rendje az adott helyen helyreállt (ezt a folyamatot nevezik *rekombinációnak*), viszont a most elszabadult szabad elektron helyén új lyuk keletkezett. Ha valamelyik szomszédos atomból elszabaduló elektron most ebbe a lyukba ugrik be, tehát rekombinálódik, a lyuk a most elszabadult atom helyére »vándorol át«.” Forrás: Internet – [http://www.puskas.hu/r\\_tanfolyam/felvezetok.pdf](http://www.puskas.hu/r_tanfolyam/felvezetok.pdf), 2. oldal

Tekintsünk egy elektromos áramot, mint különálló elektromos töltéshordozók áramlását. Ha a töltéshordozók egymástól függetlenül mozognak, akkor az áram ingadozása megadható a  $I_{Noise}(rms) = \sqrt{2eI_{dc}B}$  összefüggéssel, ahol  $e = 1,6 \cdot 10^{-19}$  coulomb (egy elektron töltése), és  $B$  pedig a mérhető sávszélesség. Például egy  $I_{dc}$  1 amper „állandó” áram, és 1 MHz sávszélesség mellett, az effektív hullámzás 0,57 mikroamper lesz. Az iménti összefüggés feltételezi, hogy egymástól független töltéshordozók alkotják az áramot. Abban az eset mikor a töltéshordozóknak egy akadályon kell áthaladniuk, mint például egy félvezető dióda p–n átmenete, a töltéshordozók nem együtt mozognak; viszont nem igaz ez a fém vezetőkre, ahol a töltéshordozókat nagyfokú kölcsönhatás jellemzi.

### 2.1.8. Elektromágneses zaj

Virtuálisan minden elektromos eszköz, mely energiát állít elő, energiát fogyaszt vagy energiát szállít, elektromágneses zajok potenciális forrása. Az ilyen elektromágneses zaj interferenciának tekinthető más rendszerek szempontjából. Általában minél nagyobb a feszültség vagy az áram és minél közelebb helyezkedik el az áramkörhöz vagy elektromos eszközhöz, annál nagyobb lesz az áramkörben tapasztalható elektromágneses zaj mértéke. Az elektromágneses zajok általános forrásai a transzformátorok, rádió és televízió adóállomások, mobiltelefonok, mikrohullámú adóállomások, váltakozó áramú vezetékek, motorok és motorindítók, generátorok, relék, oszcillátorok, fénycsövek, és villámlások (zivatarok).

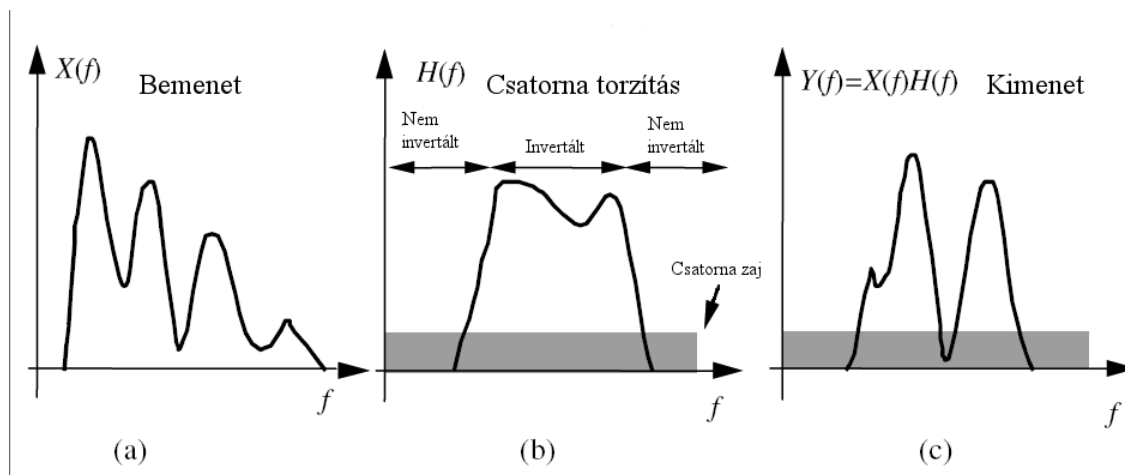
Az elektromágneses zajok két alapvető csoportba sorolhatók: elektrosztatikus zajok, és mágneses zajok. Ez két típus alapvetően különbözik egymástól, ezért más-más zajvédelem szükséges hozzájuk. Sajnos az imént felsorolt zajforrások legtöbbje a két típus kombinációját állítja elő, mely tovább bonyolítja a zajcsökkentési problémát.

Az elektrosztatikus tereket a feszültség jelenléte generálja, árammal, vagy anélkül. Az elektrosztatikus zajok egyik legáltalánosabb forrásai a fénycsövek. Mágneses terek keletkeznek elektromos áram hatására, vagy állandó mágnes jelenléte esetén. A motorok és a transzformátorok szolgálhatnak például az iménti felsorolásból, és hogy ne menjünk messzire a Föld mágneses tere is itt van környezetünkben. Ahhoz, hogy zajfeszültség indukálódjon egy vezetőben, a mágneses erővonalakat metszenie kell. A generátorok működési elve is ezen a jelenségen alapul. A váltakozó mágneses tér jelenléte – amilyen például az 50/60 Hz-es elektromos vezetékeket körülveszi – feszültséget indukál minden álló vezetőben, ahogy a mág-

neses tér kialakul, majd eltűnik (váltakozó mágneses tér). Ehhez hasonlóan, ha egy vezető mozog a Föld mágneses terében, zajfeszültség indukálódik benne, ahogy mozgás közben metszi a Föld mágneses erővonalait.

### 2.1.9. Csatornatorzítások

Ha egy jelet egy csatornán keresztül továbbítunk, a jel alakja változik, torzul a csatorna frekvenciaátvivő képességének és csillapítási paramétereinek megfelelően. A csatornatorzítás két fő megnyilvánulási formája a folytonossági torzítás és a fázistorzítás. A rádiókommunikációban létezik egy multi-út hatás is, mely szerint ha a jel több átjátszón megy keresztül, bizonyos esetekben különböző utakon haladva, ezzel a jel többféle változata érkezik a vevőbe különböző késleltetéssel és csillapítással. A csatornatorzítások csökkenthetők, vagy egészen egyszerűen szétválaszthatók a kommunikációs folyamatokban, ezért a csatorna modellezése és ki-

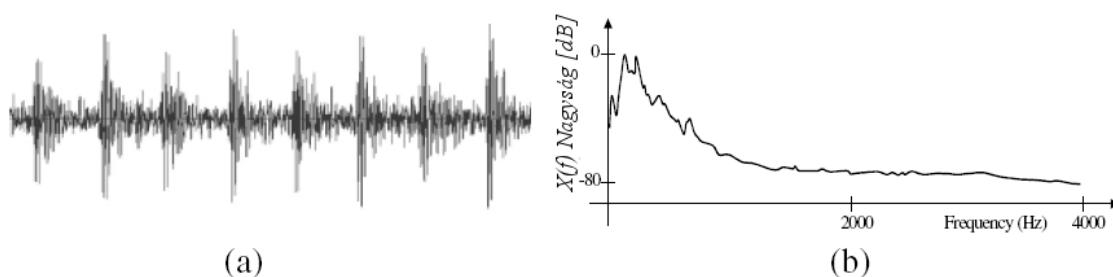


7. ábra Csatornatorzítás

egyenlítése nagyon jelentős összetevői egy modern digitális kommunikációs rendszer kialakításának. A csatorna kiegyenlítése különösen fontos a modern sejtes felépítésű kommunikációs rendszerekben mivel a csatorna karakterisztikák és terjedési paramétereik különbsége igen változatos, sokkal változatosabb, mint a vezetékes rendszerekben. A 7. ábra egy csatorna sávátviteli jellemzőit mutatja, mely két nem megfordítható és egy megfordítható régióra oszlik. A megfordítható csatornarégióban a jelek erősen torzultak elvesztek a csatornazaj miatt. A nem megfordítható csatornarégióban a jelek torzultak ugyan, visszanyerhetők. Ez a példa remekül mutatja, hogy a csatornafordító szűrőket nagy gondossággal kell elkészíteni, a nem kívánt eredmények elkerülése érdekében, mint például a zajerősítés az alacsony jel-zaj arányú frekvenciákon.

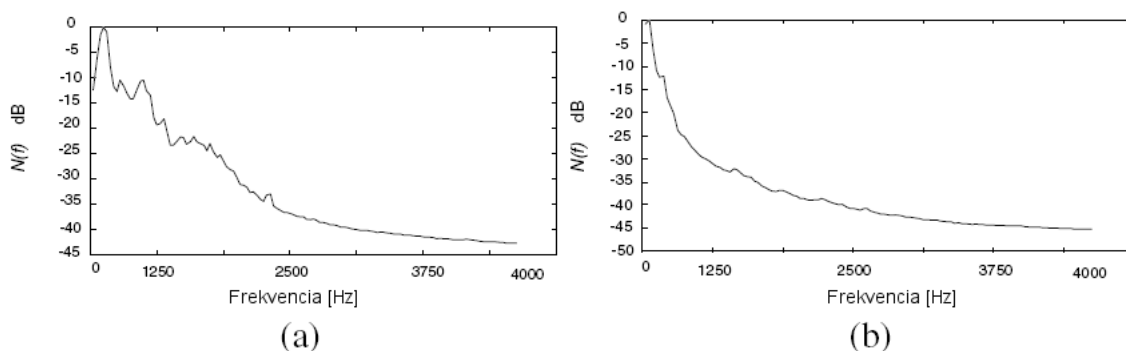
## 2.1.10. Zajmodellezés

A cél a modellezés esetén, hogy a jelben vagy zajban jelalakokat és mintákat jellemezzünk. A zaj pontos modellezéséhez egy olyan struktúrára van szükségünk, mely mind időben mind spektrumában jellemzi a zajt. A pontos zajstatisztika modellezéshez a kulcs a zajos jel megfelelő osztályozása és a hasznos jel erősítése. Ezen látszólag egyszerű jel/zaj osztályozási munka döntően függ a rendelkezésre álló hasznos jel mennyiségétől, erősségétől és a zajmodellektől, és ezen modellek alkalmazásától.



9. ábra A fúrózaj (a) idő-hullámformája (b) frekvenciaspektruma

A leghasználhatóbb szinte nélkülözhetetlen módszer az összes közül a zaj felépítésében való éleslátásra a Fourier transzformáció a zaj frekvenciaanalizálására. A 8. ábra egy elektromos fúró zaját mutatja, és ahogy az várható a zaj periodikus. Az ábra (b) része mutatja a fúró zajának spektrumát, melyen látható, hogy a zaj energiájának jelentős része a spektrum alacsonyabb frekvenciáira koncentrálódik. Tényszerűen igaz, hogy a legtöbb hangfrekvenciás jelnek és zajnak van túlnyomórészt alacsonyfrekvenciás a spektruma. Bár meg kell jegyezni, hogy a viszonylag kisebb energiájú magasabb frekvenciájú hangjelek fontos szerepet játszanak az élmény és a minőség átadásában. A 9. ábra (a) és (b) része példával szolgál az autók belsejéből származó zaj spektrumából (egy BMW és egy Volvo zajai láthatók). A



8. ábra Autók zajának spektuma (a) BMW 112 km/h, (b) Volvo 112 km/h

zajok nem statikusak, változatosak, és a következő forrásokból származhatnak:

- a.) kvázi periodikus zaj a motorból, és a jármű forgó alkatrészeiből,
- b.) zaj a kerekek és a talaj érintkezéséből és az út felületéből adódóan,
- c.) az autó körüli légáramlatok a légjáratokon, ablakokon, napfénytetőn keresztülhaladva,
- d.) az éppen előzőtt jármű zaja

Az autó zajainak jellemzői változnak a sebességgel, az út felületének állapotváltozásával, az időjárással és az autó környezetével.

A zajmodellezés legegyszerűbb módja, melyet gyakran használunk a jelenlegi gyakorlatban, hogy kiszámítjuk a zaj statisztikáját jelmentes időszakok alapján. Optimális esetben a Bayes féle jelfeldolgozási módszerekben egy készlet valószínűségi modell van előkészítve a jel- és zajfeldolgozásra. Ezek a modellek alapjául szolgálnak a jelek és zajok dekódolásának, zajos jelek felismerésének, erősítésének.

#### 2.1.10.1. Hozzáadódó fehér normális eloszlású zajmodell (AWGN<sup>1</sup>)

A kommunikációelméletben a gyakran feltételezzük, hogy a zaj hozzáadódó fehér normális eloszlást követő. Bár számos probléma esetében ez valós feltevés, és matematikailag kényelmes és használható megoldáshoz vezet, a gyakorlatban a zaj gyakran időben változó, kölcsönösen összefüggő, és nem normális eloszlást követő. Ez részben igaz az impulzív típusú zajokra, melyek nem állandóak, nem normális eloszlásúak, ezért nem modellezhetők az AWGN feltevés használatával. A nem állandó és nem normális eloszlású zajok modellezésére a Markov-féle sorozat alkalmas, stacionárius alfolyamatok leírásával.

#### 2.1.10.2. Rejtett Markov modell

A legtöbb zaj nem állandó, ez érvényes a zaj statisztikai paramétereire, úgy mint a zaj középértéke, szórásnégyzete, erőssége és spektruma, és ezek időbeli változása. A nem stacionárius zajok rejtett Markov modellek (HMM<sup>2</sup>) segítségével modellezhetők. Egy HMM lényegében stacionárius alfolyamatok véges sok állapotú Markov sorozata. Az állandó zaj modellezhető egy egyállapotú HMM segítségével. A nem állandó zaj modellezésére a többállapotú HMM alkalmas, amely képes a zaj időbeli változásának követésére, véges sok állandó állapottal. Nem normális eloszlású zajok esetében egy Gauss sűrűségi modell keverékét alkalmazhatjuk, a zaj terének modellezésére minden állapotban. Általában az állapotok száma modellen-

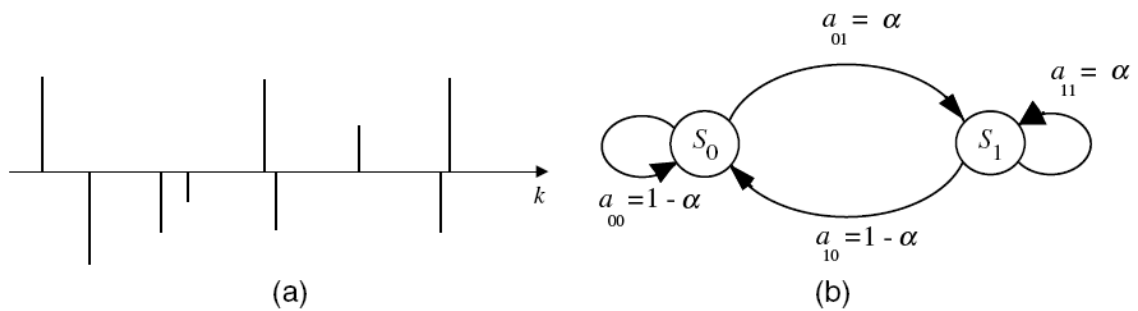
---

<sup>1</sup> AWGN: Additive White Gaussian Noise Model

<sup>2</sup> HMM: Hidden Markov Model

ként, és a keverékek száma állapotonként szükséges a zaj pontos modellezéséhez, mely a nem állandó zaj karakterisztikájától függ.

Egy példa a nem állandó zajra az impulzív zaj, melyet a 10. ábra (a) része szemléltet. Az ábra (b) részén az impulzív zaj szekvencia egy kétállapotú HMM modellje látható, ahol az  $S_0$  állapot jelképezi az „impulzus-ki” periódusokat az impulzusok között, és az  $S_1$  állapot jelképezi magát az impulzust. Azokban az esetekben ahol minden impulzus jól definiált átmeneti felépítéssel rendelkezik, előnyös lehet a többállapotú HMM magának az impulzusnak a modellezésére.



10. ábra (a) Impulzív zajszekvencia, (b) az impulzív zaj kétállapotú modellje

## 2.2. Zajmodellek

### 2.2.1. Valószínűségi modellek

A valószínűségi modellek alapja az információelmélet. Az információ maga is mennyiségileg a valószínűség logaritmusaként fejezhető ki. A valószínűségi modellek általában azzal foglalkoznak, hogy történések előfordulását megjósolják és jellemzésük véletlenszerű események alapján, különféle területeken, mint például telefonhívások számának megjóslása egy fővonalon a nap egy meghatározott időszakában, közúti forgalommodellezés, időjárás előrejelzés, pénzügyi adatmodellezés, előre jelezni a drogok hatását megadott adatok alapján az orvosi gyakorlatban. A jelfeldolgozásban a valószínűségi modelleket általában véletlenszerű jelek ingadozásának leírására alkalmazhatjuk, mint például mintafelismerés, jelkódolás és jelkiértékelés.

#### 2.2.1.1. Véletlenszerű jelek és sztochasztikus folyamatok

A jelek egyik legalapvetőbb jellemzőjük szempontjából két nagy csoportba osztályozhatók: determinisztikus (pontosan meghatározott) jelek és véletlenszerű jelek. Az időben véletlenszerű jelekre gyakran úgy hivatkozunk, hogy sztochasztikus jelek. Minden osztályban a jelek lehetnek folyamatosak és diszkrét az időben, és rendelkezhetnek folytonos értékű és diszkrét értékű amplitúdóval.

A determinisztikus jel definiálható úgy, hogy egyike az előre meghatározott görbéjű jeleknek időben és térben. A determinisztikus jelek pontos hullámszáma teljes mértékben leírható az idő függvényében, és a jel pontos értéke bármely időpillanatra kiszámítható a függvényi leírásból, és a korábbi értékek is kiszámíthatók ezen a módon. Például egy szinusz hullám  $x(t)$  modellezhető és pontosan kiszámítható mind egy másodrendű lineáris prediktív modellből, vagy a közismertebb  $x(t) = A \cdot \sin(2\pi ft + \phi)$  egyenlettel.

A véletlenszerű jeleknek nem kiszámítható hullámszámai vannak, ezért nem lehet olyan egyenletet felírni, mellyel kiszámítható lenne a jel pontos jövőbeli értéke a korábbi értékeiből. A legtöbb jel, mint például a beszéd, vagy zaj legalább részben véletlenszerű. A véletlenszerűség elmélete az információ és a zaj elméletével teljesen összekapcsolódott. Valóban, a véletlenszerű jelek feldolgozásakor a munka nagy részét jelenti az információ kinyerése a zajosan vett jelből. Ha egy jel alkalmas információhordozásra, akkor valamilyen fokú véletlenszerűségnek kell benne lennie: a kiszámítható jelek nem hordoznak információt. Ezért a jel véletlenszerű



része lehet a jel információtartalma, vagy zaj, vagy a kettő keveréke, tehát információ és zaj. Bár a véletlenszerű jelek nem teljesen kiszámíthatók, mégis gyakran mutatnak egy sor olyan statisztikailag jellemző értéket, mint például a maximum, a minimum, a középérték, a felező, a szórásnégyzet, és a teljesítményspektrum. A véletlenszerű jel leírható a statisztikája szempontjából, és a legteljesebben a valószínűségi modell szempontjából, melyből az összes statisztikája számítható.

#### 2.2.1.1.1. Sztochasztikus folyamatok

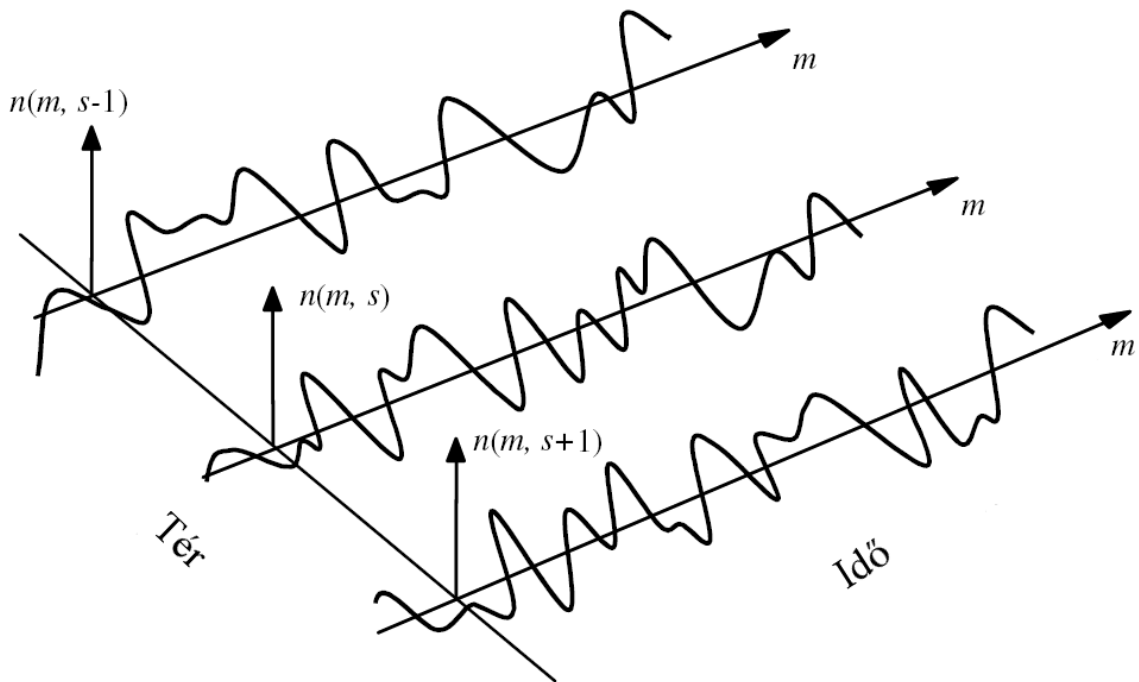
A „sztochasztikus eljárás” szakkifejezést általában a véletlenszerű folyamatok leírására használjuk, melyek folyamatos jeleket állítanak elő, mint például az emberi beszéd, vagy zaj. A jelfeldolgozási terminológiában egy sztochasztikus eljárás egy valószínűségi modell, a véletlenszerű jelek egy osztályára, például Gauss módszer, Markov módszer, Poisson módszer. Egy klasszikus példa a sztochasztikus eljárásra a részecskék úgynevezett Brown mozgása folyadékokban. A részecskék a folyadék belsejében véletlenszerűen mozognak a folyadék részecskék „bombázása” következtében. Minden részecske véletlenszerű mozgása a sztochasztikus módszer egy egyszerű elképzelése. Minden részecske folyadékban történő mozgása segíti a megértést, vagy a módszer különféle megvalósításainak terét.

#### 2.2.1.1.2. Véletlen folyamatok együttese vagy tere

A véletlenszerű folyamatok összes elképzelésének gyűjteménye a módszerek együttese vagy tere néven ismert. Példaként tekintsünk egy véletlenszerű zajt egy telekommunikációs rendszerben ahogyan azt a 11. ábra is mutatja. A zaj minden telefonvonalon időben véletlenszerűen ingadozik, és jelölhető az  $n(m,s)$  jelöléssel, ahol  $m$  a diszkrét idő mérőszám, és  $s$  jelöli a sorindexet. A zajok gyűjteménye a különböző vonalakon a zaj együttesét (vagy terét) alkotják, melyet a  $N(m)=\{n(m,s)\}$  jelöléssel jelölünk, ahol  $n(m,s)$  az  $N(m)$  zaj realizációját jelöli az  $s$  vonalon.

#### 2.2.1.2. Valószínűségi modellek

A valószínűségi modellek a véletlenszerű jelek legteljesebb matematikai leírását adják. Egy  $m$  rögzített időpillanatban egy véletlenszerű jeltől gyűjtött minta  $\{x(m,s)\}$  egy valószínűségi változó, mely különböző értékeket vesz fel az  $s$  térben. Az alapvető különbség a véletlenszerű változó és a véletlenszerű folyamat között, hogy az utóbbi idősorozatot generál. Ezért a véletlenszerű változók modellezésére használt modellek használhatók a véletlenszerű folyamatokra is.



11. ábra A véletlen zaj három realizációja a zajok terében

A valószínűségi változó tere az összes érték gyűjteménye, vagy eredmény, amit a változó felvehet. A valószínűségi változó tere részekre osztható, néhány kritérium szerint, néhány altérre. Az altér jelértékek egy gyűjteménye, melyek egy közös tulajdonsággal rendelkeznek, mint egy csoportba összezsúfolt minta, vagy a minták gyűjteménye azok amplitúdójával az adott sávon belül. Minden alteret egy eseménynek hívunk, és az esemény bekövetkezésének valószínűsége  $A$ , és  $P(A)$  az a hányados, mely az  $A$  térből származó megfigyelések eredményeinek számát  $N_A$ -t

elosztjuk az összes megfigyelés számával:  $P(A) = \frac{N_A}{\sum_i N_i}$ . Az iménti egyenletből evi-

dens, hogy a kísérletben résztvevő összes lehetséges esemény valószínűségének összege egy.

#### 2.2.1.2.1. Valószínűségi változó eloszlása, eloszlásfüggvény

Legyen adott egy  $(\Omega, \mathbf{A}, P)$  valószínűségi mező, legyen értelmezve egy  $\xi: \Omega \rightarrow \mathbf{R}$  függvény, melyre teljesül, hogy  $\{\omega \mid \xi(\omega) = x\} \in \mathbf{A}$  minden  $x \in \mathbf{R}$  esetén. ( $\xi$  az elemi események halmazát képezi le  $\mathbf{R}$ -be, értékkészlete diszkrét értékekből áll). Ekkor  $\xi$ -t diszkrét valószínűségi változónak nevezzük.  $\xi$  értékkészlete  $\{x_1, x_2, \dots\}$  megszámlálható.  $\xi$  eloszlása:  $\{p_i = P(\xi = x_i) = P(\omega \mid \xi(\omega) = x_i) \mid i=1, 2, \dots\}$  valószínűségek halmaza, annyi elemű, ahány elemű az értékkészlet, elemei mutatják a diszkrét valószínűségi változó egyes értékeinek felvételi valószínűségét. Az értékkészlet elemei nemnegatív számok, melyek összege 1.

Két diszkrét valószínűségi változó ( $\xi$  és  $\eta$ ) *együttes eloszlásán* a következő eloszlást értjük (amennyiben  $\xi$  értékészlete  $\{x_1, x_2, \dots\}$  eloszlása  $\{P(\xi=x_k) \ (k=1, 2, \dots)\}$ ,  $\eta$  értékészlete  $\{y_1, y_2, \dots\}$  eloszlása  $\{P(\eta=y_j) \ (j=1, 2, \dots)\}$ ):  $\{P(\xi=x_k \text{ és } \eta=y_j) \ (k,j=1, 2, \dots)\}$ .

Az együttes eloszlás elemei nemnegatívak, az elemek összege 1. Az együttes eloszlásban szereplő események teljes eseményrendszer alkotnak, ezért az előző állítás igaz.

A peremeloszlás két diszkrét valószínűségi változó együttes eloszlásából meghatározható az alkotó diszkrét valószínűségi változók eloszlása. Valamelyik diszkrét valószínűségi változó eloszlása szerint összegezve az együttes eloszlást, a másik diszkrét valószínűségi változó eloszlásához jutunk.

Diszkrét valószínűségi *változók függetlenek*, ha minden  $j = 1, 2, \dots$  és minden  $k = 1, 2, \dots$  esetén  $P(\xi=x_k \text{ és } \eta=y_j)=P(\xi=x_k)P(\eta=y_j)$ , azaz együttes eloszlásuk minden eleme szétbomlik az eloszlások megfelelő elemeinek szorzatára. Két diszkrét valószínűségi változó eloszlása pontosan akkor határozza meg együttes eloszlásukat, ha függetlenek. A konstans értékű valószínűségi változó minden más diszkrét valószínűségi változótól független.

Adott valószínűségi mezőben lévő  $B_i \ (i=1, 2, \dots)$  teljes eseményrendszer esetén ha ezek között nincs 0 valószínűségű esemény, akkor tetszőleges nem 0 valószínűségű  $A$  eseményre  $P(A | B) = \frac{P(A | B_i) \cdot P(B_i)}{\sum_i (P(A | B_i) \cdot P(B_i))}$ . Ez Bayes tétele.

Egy  $\xi$  valószínűségi változó *eloszlásfüggvénye*:  $F: \mathbf{R} \rightarrow [0, 1]$ ,  $F(x)=P(\xi < x)$ . Valószínűségi változó eloszlásfüggvénye balról folytonos, monoton nemcsökkenő,  $-\infty$ -ben 0-hoz tart,  $+\infty$ -ben 1-hez tart. Bármely, a valószínűségi változó eloszlásfüggvényeire vonatkozó fenti feltételeket kielégítő függvényhez létezik olyan valószínűségi változó, melynek eloszlásfüggvénye az adott függvény. A diszkrét valószínűségi változónak is létezik eloszlásfüggvénye, ez „lépcsős” alakú.  $\xi$  valószínűségi változó esetén  $P(a \leq \xi < b) = P(\xi < b) - P(\xi < a) = F(b) - F(a)$ , ha  $a < b$ .  $\xi$  és  $\eta$  *valószínűségi változók együttes eloszlásfüggvénye*  $F(x,y)=P(\xi < x \text{ és } \eta < y) \ (x,y \in \mathbf{R})$ .  $\xi$  és  $\eta$  *valószínűségi változók függetlenek*, ha együttes eloszlásfüggvényük megegyezik eloszlásfüggvényeik szorzatával. Valószínűségi változók együttes eloszlásfüggvénye mindkét változójában monoton nemcsökkenő, balról folytonos,  $\lim_{x \rightarrow -\infty} F(x,y) = \lim_{y \rightarrow -\infty} F(x,y) = 0$ ,  $\lim_{y \rightarrow \infty, x \rightarrow \infty} F(x,y) = 1$ ; ha  $a_1 < b_1$  és  $a_2 < b_2$ , akkor  $F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) - F(a_1, a_2) \geq 0$ . Bármely, a valószínűségi változók együttes eloszlásfüggvényére vonatkozó ezen feltételeket kielégítő függvényhez létezik két olyan valószínűségi változó, melyek-

nek együttes eloszlásfüggvénye az adott függvény. Az utolsó feltétel  $P(\xi \in [a_1, b_1] \text{ és } \eta \in [a_2, b_2]) \geq 0$ -t jelenti.

### 2.2.1.2.2. Sűrűségfüggvény

Egy adott eseménytéren értelmezett  $\xi: \Omega \rightarrow \mathbf{R}$  függvény *valószínűségi változó*, ha  $\{\omega \mid \xi(\omega) < x\} \in \mathbf{A}$  bármely  $x \in \mathbf{R}$  esetén. A diszkrét valószínűségi változó teljesíti az iménti definíció követelményeit. A definíció a diszkrét valószínűségi változók általánosítása: folytonos esetben  $P(\xi=x)=0$  minden  $x \in \mathbf{R}$  esetén. Valószínűségi változót (ha nem hangsúlyozottan diszkrét) a definíció szerinti értelemben kell érteni.

Egy  $\xi$  valószínűségi változó *folytonos*, ha eloszlásfüggvénye abszolút folytonos. Folytonos valószínűségi változó esetén  $P(\xi=x)=0$  bármely  $x \in \mathbf{R}$  esetén.

Belátható, hogy  $0 \leq P(\xi=x) = \lim_{\Delta x \rightarrow 0} P(x \leq \xi < x + \Delta x) = \lim_{\Delta x \rightarrow 0} (F(x + \Delta x) - F(x)) \leq \lim_{\Delta x \rightarrow 0} ((F(x) + K\Delta x) - F(x)) = \lim_{\Delta x \rightarrow 0} (K\Delta x) = 0$ .

Bármely  $\xi$  valószínűségi változó esetén bármely  $x \in \mathbf{R}$ -re  $P(a \leq \xi \leq b) = F(b+0) - F(a)$ ,  $P(a < \xi < b) = F(b) - F(a-0)$ ,  $P(a < \xi \leq b) = F(b+0) - F(a-0)$  teljesül. Folytonos esetben ezek mind egyenlők  $P(a \leq \xi < b)$ -vel. Folytonos  $\xi$  valószínűségi változó *sűrűségfüggvénye*

$f: \mathbf{R} \rightarrow \mathbf{R}$ , ha  $\int_{-\infty}^x f(t) dt = F(x)$  teljesül minden  $x \in \mathbf{R}$  esetén. Folytonos valószínűségi

változó sűrűségfüggvényére és eloszlásfüggvényére teljesül az  $F'(x)=f(x)$  összefüggés. Diszkrét valószínűségi változónak nincs sűrűségfüggvénye.

A sűrűségfüggvény nemnegatív, az egész számegegyenesen vett integrálja 1. Bármely, a valószínűségi változó sűrűségfüggvényére vonatkozó, ezen feltételeket kielégítő függvényhez létezik olyan valószínűségi változó, melynek sűrűségfüggvénye az adott függvény. Diszkrét esetben hasonló feltétel a  $p_i \geq 0$ ,  $\sum p_i = 1$  tulajdonságok: az eloszlás játssza a sűrűségfüggvény szerepét is.

$\xi$  és  $\eta$  abszolút folytonos *valószínűségi változók együttes sűrűségfüggvénye*

$f(u, v)$ , ha rá teljesül az  $\int_{-\infty}^x \int_{-\infty}^y f(u, v) dv du = F_{\xi, \eta}(x, y)$  egyenlőség.

Az együttes sűrűségfüggvény megegyezik az együttes eloszlásfüggvény második parciális deriváltjával (először az első, majd a második változó szerint deriválva). Abszolút folytonos  $\xi$  és  $\eta$  valószínűségi változók esetén függetlenségük egyenértékű  $f_{\xi, \eta}(u, v) = f_{\xi}(u)f_{\eta}(v)$  teljesülésével. Az együttes sűrűségfüggvény nemnegatív, az egész valós számsíkon vett integrálja 1. Az együttes sűrűségfüggvényből meghatározható az alkotó valószínűségi változók sűrűségfüggvénye. Az együttes sűrűség-

függvényt az egyik változó szerint az egész valós számegyenesen integrálva a másik valószínűségi változó sűrűségfüggvényét kapjuk.

### 2.2.1.3. Állandó és nem állandó véletlen folyamatok

Bár a jel  $x(m)$ -mel jelölt amplitúdója  $m$  periódussal hullámzik, a jel által generált folyamat jellemzője lehet időben állandó vagy időben változó. Nem állandó folyamat például a beszéd, melynek hangossága és spektrális jellemzői, összetevői folyamatosan változnak, ahogyan a beszélő különféle hangokat képez. A folyamat állandónak tekinthető, ha a valószínűségi modellbeli jellemzői idő invariánsak, ellenkező esetben nem állandó – mindkettőt a 12. ábra szemlélteti. Az állandó tulajdonság magába foglalja, hogy a jel összes paramétere, úgy mint középérték, variancia, energiaspektrum összetétele, és a folyamat magasabb rendű momentumai idő invariánsak. A gyakorlatban az állandóságnak különféle fokai ismeretesek. Van olyan, hogy a statisztikák egy halmaza állandó, a többi pedig nem. Például lehet egy véletlen folyamatnak idő invariáns középértéke, de az idő függvényében változó energiája.



12. ábra Kvázistacionárius és nemstacionárius beszédsezmensek

### 2.2.1.3.1. Szigorú értelemben vett állandó folyamatok

Egy  $X(m)$  véletlenszerű folyamat szigorú értelemben véve stacionárius, ha minden megoszlása, és statisztikai paramétere idő invariáns. A szigorú értelemben vett állandóság magában foglalja, hogy az  $n$ -edik rendű eloszlása transláció invariáns minden  $n = 1, 2, \dots$ , vagyis:

$$P[x(m_1) \leq x_1, x(m_2) \leq x_2, \dots, x(m_n) \leq x_n] = P[x(m_1 + \tau) \leq x_1, x(m_2 + \tau) \leq x_2, \dots, x(m_n + \tau) \leq x_n]$$

Az egyenletből a szigorú értelemben vett állandó folyamat statisztikája beleértve a középértéket, a korrelációt és az energiaspektrumot, idő invariáns lesz, ezért:  $E[x(m)] = \mu_x$ ,  $E[x(m)x(m+k)] = r_{xx}(k)$  és  $E[|X(f, m)|^2] = E[|X(f)|^2] = P_{xx}(f)$ , ahol  $\mu_{xx}$ ,  $r_{xx}(m)$  és  $P_{xx}(f)$  rendre az  $x(m)$  jel középértéke, az autokorrelációja és az energiaspektruma, és  $X(f, m)$  jelöli  $x(m)$  frekvencia-idő spektrumát.

### 2.2.1.3.2. Nem szigorú értelemben vett állandó folyamatok

Míg a szigorú értelemben vett állandó (stacionárius) folyamatok esetében a minden statisztikának idő-invariánsnak kell lennie, a nem szigorú értelemben vett stacionárius folyamatokra nem vonatkozik az összes korlátozás. Egy ilyen folyamat középértéke és az autokorrelációs függvénye idő invariáns:  $E[x(m)] = \mu_x$ ,  $E[x(m)x(m+k)] = r_{xx}(k)$ . A két folyamat definíciójából kiderül, hogy minden szigorú értelemben vett stacionárius folyamat nem szigorú értelemben is az, viszont ennek a fordítottja nem igaz.

### 2.2.1.3.3. Nem stacionárius folyamatok

Egy véletlenszerű folyamat nem stacionárius, ha eloszlása vagy statisztikája időben változik. A legtöbb sztochasztikus folyamat, mint például a videojelek, audiojelek, gazdasági adatok, meteorológiai adatok, stb., nem stacionárius, mert egy-egy olyan rendszer generálja őket, melynek környezeti paraméterei időben változnak. Például a beszéd nem stacionárius, mert egy időben változó artikulációs rendszer állítja elő. A beszéd hangerőssége, frekvencia-összetétele időben változó, mely változás néha teljesen váratlan. Az időben változó folyamatok modellezhetők stacionárius véletlenszerű folyamatok egy kombinációjával.

### 2.2.1.4. Véletlen folyamatok várható értéke

A várható érték központi szerepet játszik a jelek modellezésében és feldolgozásában. Továbbá a véletlenszerű folyamatok valószínűségi modelljei általában a várható érték függvényeként kerülnek kifejezésre. Például a normális eloszlás el-

oszlásfüggvénye a folyamat középértékét és kovarianciáját tekintve egy exponenciális függvényként definiálható, a Poisson eloszlás eloszlásfüggvénye pedig a folyamat középértékével definiálható. A jelfeldolgozási alkalmazásokban mindig megvan a jel megfelelő statisztikai modellje – például a normális eloszlás eloszlásfüggvénye –, és a modell kiegészítésére szükség van a várható paraméterek értékére. Azonkívül sok jelfeldolgozási algoritmusban, mint például a spektrális kivonásban a zajcsökkentésre, vagy a lineáris előrejelzésben, lényegében amire szükségünk van az a jel középértékének vagy a korrelációs függvényének kiszámítása. Egy véletlenszerű folyamat  $h(X(m_1), X(m_2), \dots, X(m_M))$  függvényének várható értékének kiszámítása a következőképpen definiálható:

$$E[h(X(m_1), \dots, X(m_M))] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} h(x_1, \dots, x_M) f_{X(m_1) \dots X(m_M)}(x_1, \dots, x_M) dx_1 \dots dx_M$$

A legfontosabb és széles körben használt várható értékek a középérték, a korreláció, a kovariancia és a energiaspektrum.

#### 2.2.1.4.1. A középérték

A egy jelsorozat középértéke nagyon fontos szerepet játszik a jelfeldolgozásban és paraméterkiszámításban a zajos jelek esetében. Például egy zajos jel esetében az optimális lineáris kiszámítási módszer az interpoláció a középérték és a megfigyelt jel zajjal terhelt értéke között. Egy  $[X(m_1), \dots, X(m_M)]$  véletlenszerű vektor középértéke alatt értjük a teljes jelfolyam alatt mutatott átlagértéket, mely a következőképpen definiálható:

$$E[X(m_1), \dots, X(m_M)] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (x_1, \dots, x_M) f_{X(m_1) \dots X(m_M)}(x_1, \dots, x_M) dx_1 \dots dx_M$$

#### 2.2.1.4.2. Autokorreláció

A korrelációs függvény és annak Fourier transzformációja, az energiaspektrum sűrűsége minták és struktúrák modellezésére és felismerésére használatos a jelfeldolgozások során. A kölcsönösségi együtthatók központi szerepet játszanak a jelfeldolgozásban, a telekommunikációs rendszerekben, beleértve a prediktív kódolókat, kiegyenlítőket, digitális dekódereket, késleltetés kiszámítókat, osztályozókat és jelvisszaállító rendszereket. Egy  $X(m)$  véletlenszerű folyamat autokorrelációs függvénye, melyet  $r_{xx}(m_1, m_2)$ -vel jelölünk, a következőképpen definiálható:

$$r_{xx}(m_1, m_2) = E[x(m_1)x(m_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(m_1)x(m_2) f_{X(m_1), X(m_2)}(x(m_1), x(m_2)) dx(m_1) dx(m_2)$$

Az  $r_{xx}(m_1, m_2)$ -vel jelölt autokorrelációs függvény az  $X$  folyamat végeredményének  $m_1$  és  $m_2$  időpillanatbeli hasonlóságának mértékét vagy a kölcsönösségi relációját mutatja meg. Ha a véletlenszerű folyamat végeredménye az  $m_1$  és az  $m_2$  időpillanatbeli értéke nincs hordoz relációt, akkor  $X(m_1)$  és  $X(m_2)$  függetlennek vagy korrelálatlannak mondható, tehát  $r_{xx}(m_1, m_2) = 0$ . A nem szigorú értelemben vett stacionárius folyamatok esetében az autokorrelációs függvény idő invariáns és a két időpillanat között eltelt idő hosszának függvénye, vagyis  $m = m_1 - m_2$ :  $r_{xx}(m_1 + \tau, m_2 + \tau) = r_{xx}(m_1, m_2) = r_{xx}(m_1 - m_2) = r_{xx}(m)$ . Egy valós értékű nem szigorú értelemben vett stacionárius folyamat autokorrelációs függvénye szimmetrikus a következő tulajdonságokkal:  $r_{xx}(-m) = r_{xx}(m)$  és  $r_{xx}(m) \leq r_{xx}(0)$ . Ez esetben az  $r_{xx}(0)$  jelenti a nulla középértékű jel energiáját.

#### 2.2.1.4.3. Autokovariancia

Az autokovariancia függvény egy  $X(m)$  véletlenszerű folyamat  $c_{xx}(m_1, m_2)$  jelöléssel ellátott szórásának vagy szóródásának mérésére szolgál a folyamat középértéke körül és a következőképpen definiálható:

$$c_{xx}(m_1, m_2) = E[(x(m_1) - \mu_x(m_1))(x(m_2) - \mu_x(m_2))] = r_{xx}(m_1, m_2) - \mu_x(m_1)\mu_x(m_2),$$

ahol a  $\mu_x(m)$  jelenti az  $X(m)$  véletlenszerű folyamat középértékét. Nulla középértékű folyamatok esetében az autokovariancia és az autokorreláció függvények azonosak. A  $c_{xx}(m_1, m_2)$  a folyamat varianciáját (szórásnégyzetét) jelenti. Stacionárius folyamatok esetében az autokovariancia függvény az előző egyenletből a következőképpen alakul:  $c_{xx}(m_1, m_2) = c_{xx}(m_1 - m_2) = r_{xx}(m_1 - m_2) - \mu_x^2$ .

#### 2.2.1.4.4. Spektrális energiasűrűség

Egy véletlenszerű folyamat energiájának spektrális sűrűségfüggvénye, más néven energiaspektruma megadja a jel egyes frekvenciákhoz tartozó energiáinak eloszlását a teljes spektrumon. A nem szigorú értelemben vett stacionárius folyamat  $X(m)$  energiaspektruma a Wiener–Khinchin tétel alapján úgy definiálható, hogy az autokorrelációs függvény Fourier transzformáltja és a következőképpen írható fel:  $P_{XX}(f) = E[X(f)X^*(f)] = \sum_{m=-\infty}^{\infty} r_{xx}(m)e^{-j2\pi fm}$ , ahol  $r_{xx}(m)$  és  $P_{XX}(f)$  rendre az  $x(m)$  autokorrelációs függvénye és energiaspektruma, és  $f$  pedig a frekvenciaváltozó. Egy valós értékű stacionárius folyamat autokorrelációs függvénye szimmetrikus, és az energiaspektruma a következőképpen írható fel:

$$P_{XX}(f) = r_{xx}(0) + \sum_{m=1}^{\infty} 2r_{xx}(m) \cos(2\pi fm)$$



A spektrális sűrűségfüggvény egy nemnegatív valós értékészletű felvevő függvény, megadása pedig watt per herzben  $\left[\frac{\text{W}}{\text{Hz}}\right]$  történik. Az autokorrelációs so-

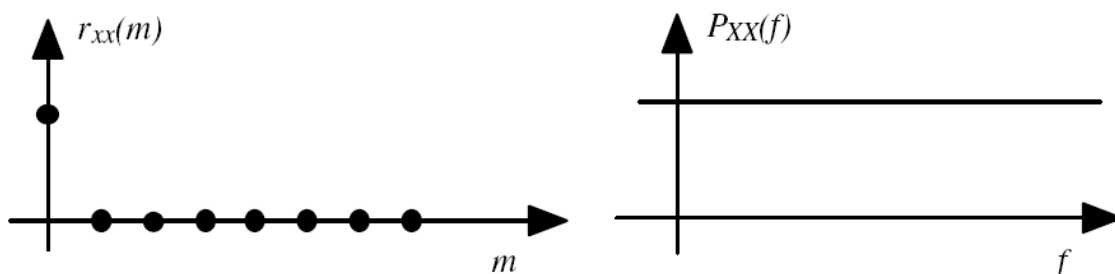
rozat a véletlenszerű folyamat energiaspektrumából az iménti egyenlet felhasználásával kapható meg az inverz Fourier transzformáció alkalmazásával a követke-

zőképpen:  $r_{xx}(m) = \int_{-\frac{1}{2}}^{\frac{1}{2}} P_{XX}(f) e^{j2\pi fm} df$ . Az autokorreláció és az energiaspektrum a

folyamat másodrendű statisztikai, rendre időben és frekvenciatartományban.

#### Példa – fehér zaj

A következő ábra a fehér zaj energiaspektrumát és autokorrelációját mutatja.



13. ábra A fehér zaj autokorrelációja és energiaspektruma

Egy olyan zajt, amely korrelálatlan, egymástól független mintákból áll, fehér zajnak nevezünk. Egy stacionárius fehér zaj  $n(m)$  autokorrelációja definiálható:

$r_{nn}(k) = E[n(m)n(m+k)] = \begin{cases} \text{Zajenergia} & k = 0 \\ 0 & k \neq 0 \end{cases}$ . Ez az egyenlet egy korrelálatlan fehér

zaj definíciója szerinti matematikai állítás. A frekvenciatartományra ekvivalens leírás adható meg, az  $r_{nn}(k)$  Fourier transzformációjával a következőképpen:

$P_{NN}(f) = \sum_{k=-\infty}^{\infty} r_{nn}(k) e^{-j2\pi fk} = r_{nn}(0) = \text{zajenergia}$ . A stacionárius fehér zaj energiaspekt-

ruma azonosan terjed szét a sáv teljes szélességében és időben. A fehér zaj az egyik legnehezebben eltávolítható zajforma, mert nem határozható meg a szerkezete sem időben, sem a frekvenciasávok szerint.

#### Példa – impulzív zaj

Az impulzív zaj egy véletlenszerű kétállapotú („be/ki”) impulzussorozatból áll, véletlenszerű amplitúdóval és időben véletlenszerű előfordulással. Egy véletlenszerű impulzív zaj sorozat  $n_i(m)$  egy kétállapotú véletlenszerű amplitúdó modulált sorozatként modellezhető, a következő szerint:  $n_i(m) = n(m)b(m)$ , ahol  $b(m)$  egy

kétállapotú sorozat, mely egy impulzus jelenlétét vagy hiányát jelzi, és  $n(m)$  pedig egy véletlenszerű zaj. Tegyük fel, hogy az impulzív zaj egy korrelálatlan folyamat, így az impulzív zaj, mint egy kétállapotú folyamat autokorrelációja a következőképpen definiálható:  $r_m(k, m) = E[n_i(m)n_i(m+k)] = \sigma_n^2 \delta(k)b(m)$ , ahol  $\sigma_n^2$  a zaj varianciája. Ebben az egyenletben az autokorrelációs függvény egy kétállapotú függvényként lett kifejezve, mely az impulzív zaj  $m$  időpillanatbeli be/ki állapotától függ. Az impulzív zaj sorozat energiaspektruma az imént leírt autokorrelációs függvény Fourier transzformáltjaként kapható meg a következő összefüggéssel:  $P_{NN}(f, m) = \sigma_n^2 b(m)$ .

#### 2.2.1.4.5. Két véletlenszerű folyamat együttes statisztikája

A jelfeldolgozás során gyakran találkozunk azzal a problémával, például egy több érzékelőből álló rendszer esetében, hogy nem csak egy véletlenszerű folyamat van jelen. Az együttes statisztika és együttes eloszlás leírja a két vagy több folyamat kapcsolatát egymással. Két diszkrét időbeosztással mintavételezett véletlenszerű folyamat  $x(m)$  és  $y(n)$  együttes eloszlásfüggvénye a következő jelöléssel adható meg:  $f_{X(m_1)\dots X(m_M), Y(n_1)\dots Y(n_N)}(x_1, \dots, x_M, y_1, \dots, y_N)$ . Ha a két véletlenszerű folyamat  $X(m)$  és  $Y(n)$  korrelálatlanok, az együttes eloszlásfüggvényük a két eloszlásfüggvény szorzataként adható meg a következő felírással:

$$f_{X(m_1)\dots X(m_M), Y(n_1)\dots Y(n_N)}(x_1, \dots, x_M, y_1, \dots, y_N) = f_{X(m_1)\dots X(m_M)}(x_1, \dots, x_M) f_{Y(n_1)\dots Y(n_N)}(y_1, \dots, y_N)$$

#### 2.2.1.4.6. Kereszt-korreláció és kereszt-kovariancia

Két véletlenszerű folyamat kereszt-korrelációja az alábbiak szerint adható

$$\text{meg: } r_{xy}(m_1, m_2) = E[(x(m_1)y(m_2))] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(m_1)y(m_2)f_{X(m_1)Y(m_2)}(x(m_1), y(m_2))dx(m_1)dy(m_2)$$

Egy  $r_{xy}(m_1, m_2)$  nem szigorú értelemben vett stacionárius folyamat esetében a korrelációs függvény csak az  $m_1$  és  $m_2$  időpillanatok különbségétől függ, vagyis  $m = m_1 - m_2$ , tehát:  $r_{xy}(m_1 + \tau, m_2 + \tau) = r_{xy}(m_1, m_2) = r_{xy}(m_1 - m_2) = r_{xy}(m)$ .

A kereszt-kovariancia függvény a következőképpen definiálható:  $c_{xy}(m_1, m_2) = E[(x(m_1) - \mu_x(m_1))(y(m_2) - \mu_y(m_2))] = r_{xy}(m_1, m_2) - \mu_x(m_1)\mu_y(m_2)$ . Nulla középértékkel rendelkező folyamatok esetében a kereszt-korreláció és a kereszt-kovariancia függvények azonosak. Nem szigorú értelemben vett stacionárius folyamatok esetében a kereszt-kovariancia függvény az előző egyenlet alapján:  $c_{xy}(m_1, m_2) = c_{xy}(m_1 - m_2) = r_{xy}(m_1 - m_2) - \mu_x\mu_y$ .

## 2.2.1.4.7. Kereszt-energiáspektrum eloszlás és koherencia

Két véletlenszerű folyamat  $X(m)$  és  $Y(m)$  kereszt-energiáspektrum eloszlása a két folyamat kereszt-korrelációs függvényének Fourier transzformációjával definiálható:

$$P_{XY}(f) = E[X(f)Y^*(f)] = \sum_{m=-\infty}^{\infty} r_{xy}(m)e^{-j2\pi fm}.$$

Ahogy a kereszt-korreláció a kereszt-energiáspektrum eloszlása is a két jel hasonlóságának mértékét mutatja, vagy a két folyamat spektrális koherenciáját, mely a kereszt-energiáspektrum eloszlásának normalizált formulája és a következőképpen definiált:

$$C_{XY}(f) = \frac{P_{XY}(f)}{\sqrt{P_{XX}(f)P_{YY}(f)}}.$$

A koherencia függvény olyan alkalmazásokban használatos, mint az időkésettetés kiszámítása és a jel-zaj arány meghatározása (mérése).

## 2.2.2. A rejtett Markov modell

A rejtett Markov modelleket a véletlenszerű nem stacionárius jelek statisztikai modellezésére használhatjuk hatékonyan. Az ilyen jelek csoportjába tartozik az emberi beszéd vagy az időben változó zaj. A HMM egy Markov láncot ad meg, mely az időbeli változást vagy a terek változását állapotfüggő stacionárius alfolyamatokra bontja. Egy HMM lényegében egy véges állapothalmazzal rendelkező Bayes folyamat, Markov előzménnyel, az egyes állapotok között átmenetek modellezésére, és egy halmaz valószínűségi eloszlásfüggvény a jel véletlenszerű variációinak modellezésére minden állapoton belül.

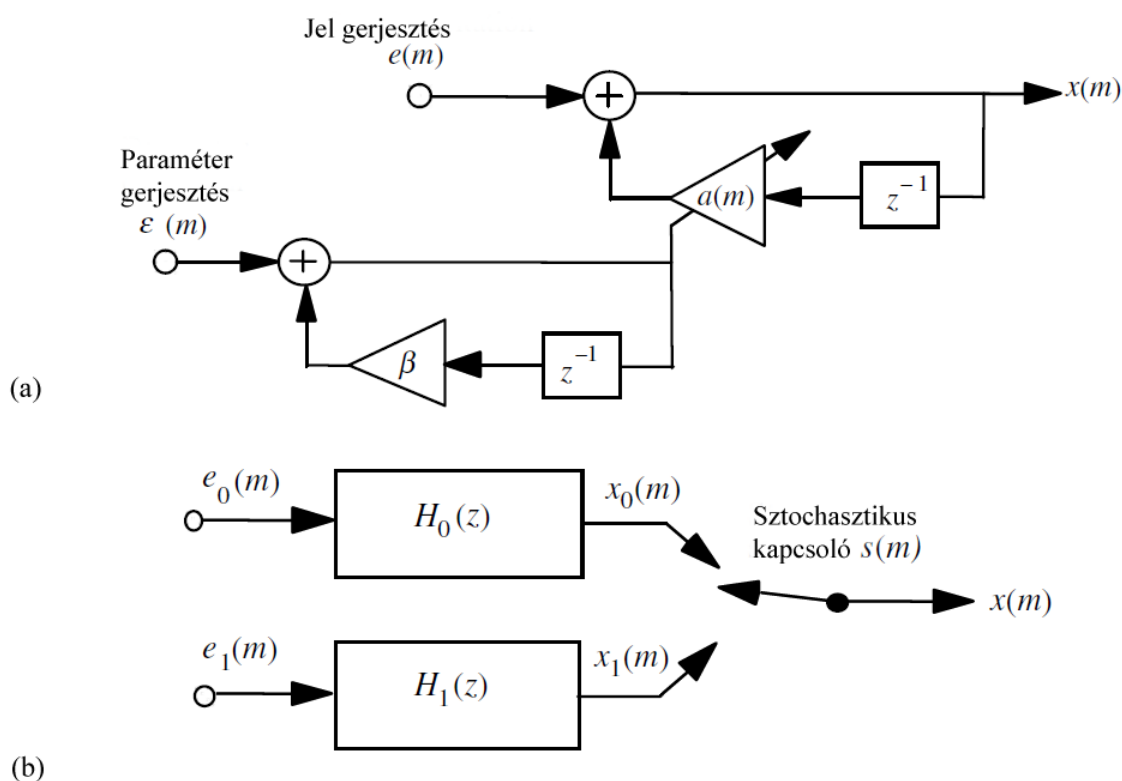
2.2.2.1. Nem stacionárius folyamatok statisztikai modellje

Egy nem stacionárius folyamat definiálható oly módon, hogy egy olyan folyamat, melynek statisztikai paraméterei az idővel változnak. A legtöbb „természetesen előállított” jel, mint például a hangjelek, képek, biomedikai és szeizmikus jelek nem stacionárius jelek, mivel az azokat előállító rendszerek paraméterei és a környezet, melyben terjednek változnak az idő múlásával.

A nem stacionárius jelek egy kétrétegű sztochasztikus folyamatként modellezhetők, egy rejtett folyamattal az észlelhető jel statisztikai paramétereinek időbeli változásainak kontrollálására. Általában a nem stacionárius folyamatok két alapvető nagy csoportba sorolhatók:

- a.) folytonosan változó állapotú folyamatok
- b.) véges sok állapotú folyamatok

A folytonosan változó állapotú folyamatok definíciós szerint az időben folytonosan változó jelek statisztikájának alapjául szolgálnak. Ilyenek például az audiojelek, a beszéd és a zene, melyeknek az energiája és spektrális összetétele az idővel változik. Véges sok állapottal rendelkező folyamat az olyan statisztikai jellemzőkkel rendelkező folyamat, mely képes váltani a véges sok stacionárius vagy nem stacionárius állapot között. Például az impulzív zaj egy ilyen kétállapotú folyamat. Folytonosan változó folyamatok közelíthetők hozzájuk illő véges sok állapotú folyamattal. A következő ábra egy nem stacionárius autoregresszív (AR) folyamatot mutat be. Ez a folyamat a jel paramétereire egy rejtett stacionárius AR modell és magára a jelre egy időben változó AR modell kombinációját alkalmazza. A



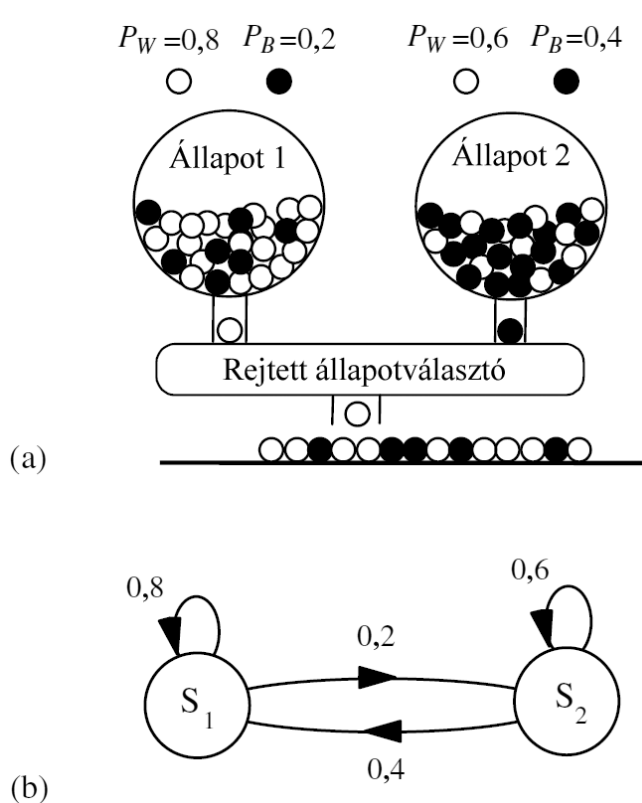
14. ábra Folytonosan változó (a) és kétállapotú (b) AR folyamat

rejtett modell vezérli a nem stacionárius AR modell időben változó paramétereit. Ehhez a modellhez a megfigyelt jel egyenlete és a paraméter állapotegyenlete a következő módon írható fel: a megfigyelt jel  $x(m) = a(m)x(m-1) + e(m)$  és a rejtett állapot  $a(m) = \beta a(m-1) + \varepsilon(m)$ , ahol  $a(m)$  a megfigyelt AR folyamat időben változó együtthatója és  $\beta$  pedig a rejtett állapotirányító folyamat együtthatója. A 14. (b) ábra egy egyszerű példát mutat a véges sok állapotú nem stacionárius folyamat kétállapotú autoregresszív megvalósítására, amelyben egy közvetlen kapcsoló választja ki minden időpillanatban a két AR modell közül az egyiket, melyet a

kimenetre csatol. Ehhez a modellhez a kimeneti jel  $x(m)$  a következő egyenlettel fejthető ki:  $x(m) = \bar{s}(m)x_0(m) + s(m)x_1(m)$ , ahol a kétállású kapcsoló  $s(m)$  kiválasztja a folyamat állapotát  $m$  időpillanatban, és  $\bar{s}(m)$  jelöli  $s(m)$  ellentettjét.

### 2.2.2.2. Rejtett Markov modellek

A rejtett Markov modell (HMM) egy kétrétegű véges sok állapotú folyamat, egy rejtett Markov folyamattal, mely irányítja az állapotok kiválasztását a figyelembe veendő folyamatban. Mint egy egyszerű illusztrációt figyeljük meg a következő ábrán látható kétállapotú Markov folyamatot, mely két tartályt mutat különböző keverékű fehér és fekete színű golyókkal. A fekete és fehér golyók előfordulásának valószínűségét jelöljük rendre  $P_B$  és  $P_W$  jelölésekkel, ahogyan az ábra mutatja. Tegyük fel, hogy folytonos idő intervallumban egy rejtett kiválasztó folyamat kivá-



15. ábra Kétrétegű véletlenszerű folyamat (a), és rejtett Markov modellje (b)

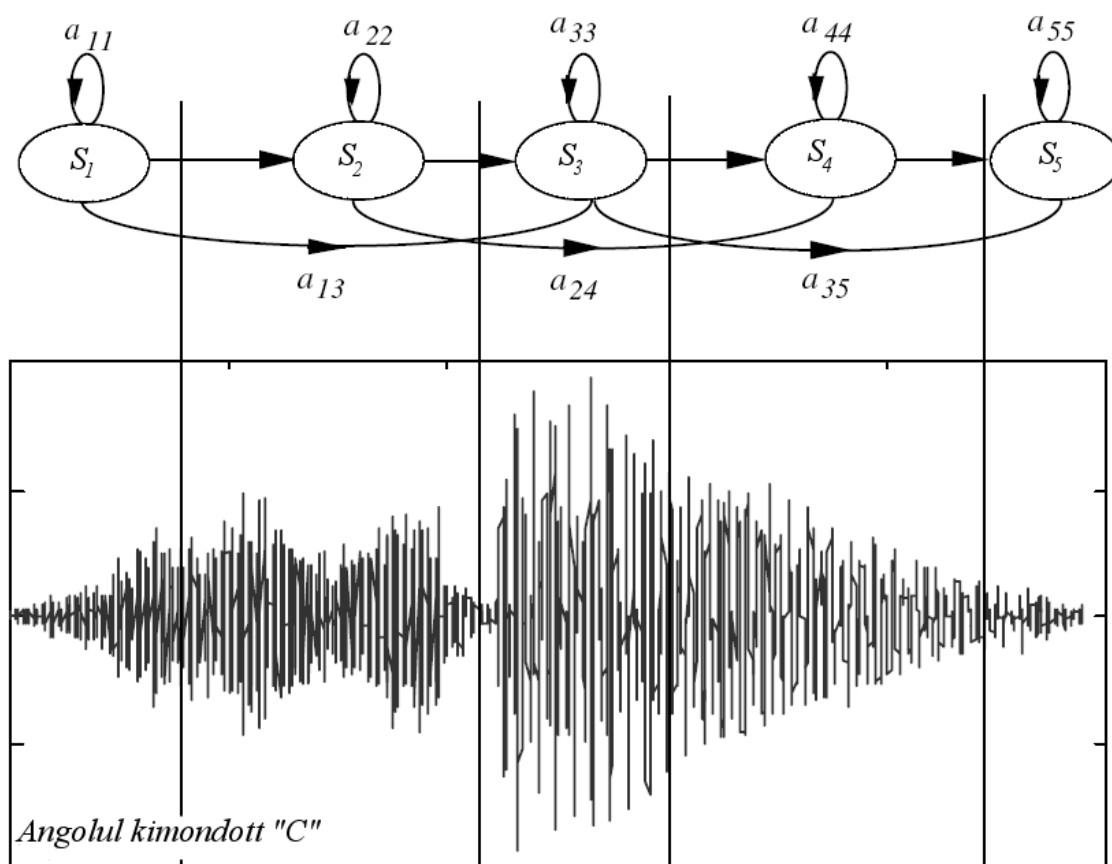
lasztja az egyik tartályt, hogy kiejtsen egy golyót. A golyót kiejtése után pótoljuk, tehát a fehér és fekete golyók számára nincs hatása a műveletnek. Minden tartályt úgy tekinthetünk, hogy az egy alapvető állapota a kimeneti folyamatnak. Most tegyük fel, hogy a rejtett tartály-kiválasztót a következő szabály irányítja: bármely időpillanatban, ha a kiválasztott tartályból fehér golyót kapunk, akkor nem választunk másik tartályt, ha feketét akkor a másik tartályt választjuk. Ez egy példa a Markov folyamatra, mert a folyamat következő állapota függ a jelenlegi állapottól, ahogyan azt az ábra kétállapotú modellje is mutatja. Ebben a modellben a megfigyelhető végeredmény nem jelzi egyértelműen az alapul vett állapotot, mivel mindkét állapot (tartály) alkalmas minkét színű (fekete és fehér) golyó kibocsátására.

Általában a Markov modellnek  $N$  állapota van, és mindegyik állapot más szegmensét modellezi a vizsgált jelnek. A rejtett Markov modell alkalmazható egy

időben változó véletlenszerű folyamat modellezésére, mint egy  $N$  elemű stacionárius vagy kvázi-stacionárius elemi alfolyamat valószínűségi Markov lánc.

### 2.2.2.2.1. A rejtett Markov modellek fizikai értelmezése

A modell fizikai értelmezéséhez nézzük meg a következő, 16. ábrát. Az ábra balról jobbra mutatja egy angolul kiejtett „C”<sup>1</sup> hang modelljét és vele együtt a hang hullámformájának térképét is. Általában kétféle eltérés van a beszéd és az egyéb sztochasztikus jelek között: a spektrális összeállítás változása, és az időbeosztás vagy artikuláció változása. Egy rejtett Markov modellben ezek a változások állapot megjegyzéssel és az egyes állapotok közötti átmenetek valószínűségével modellezhetők. Használható módszer a rejtett Markov modell interpretálására és alkalmazására, hogy a modell minden állapotát úgy tekintjük, hogy a sztochasztikus folyamat egy szegmensét modellezi. Az ábrán az  $S_1$  állapot modellezi az első szeg-



16. ábra Ötállapotú balról–jobbra HMM beszédmodell

<sup>1</sup> Az angol abc 3. betűjének, a c-nek kiejtése a magyartól jelentősen eltér, fonetikusan a következőképpen írható le: [si:]

menet, a  $S_2$  a másodikat és így tovább. Minden állapotnak kell lennie egy olyan mechanizmusának, mely összeilleszti a szegmensek véletlenszerű variációinak különböző realizációit, melyet modelleznek. Az állapotátmenetek valószínűségei egy mechanizmust biztosítanak a különböző állapotok kapcsolatára, és a variációk modellezésére az időskála jelek alatti időtartamára minden állapotban. Például ha a beszéd egy szegmensének vége kinyúlt, köszönhetően a lassú artikulációnak, hozzáigazítható lesz több önmagába visszatérő átmenettel az állapotához, mely a szegmenst modellezi. Ellenben ha a szegmens elhallgatott a gyors beszédnek köszönhetően, akkor a következő állapotra ugrással illeszt ebben az esetben. Az állapot tapasztalati eloszlásfüggvényei modellezi a jel állapotokhoz rendelt szegmenseinek spektrális összeállításának valószínűségi eloszlását.

#### 2.2.2.2.2. A rejtett Markov modell paraméterei

A rejtett Markov modellnek a következő paraméterei vannak:

- Az állapotok száma  $N$ . Ezt általában beállítjuk az egymástól különböző vagy elemi sztochasztikus események számára egy jelfolyamatban. Például egy impulzív zajhoz hasonló kétállapotú folyamat modellezése esetén  $N$  értéke 2 lesz, míg szavankénti beszéd esetében  $N$  értékét 5 és 10 közé szokás választani.
- Állapot-átmeneti valószínűségi mátrix  $\mathbf{A} = \{a_{ij}, i, j = 1, \dots, N\}$ . Ez biztosít kapcsolatot vagy Markov hálózatot hoz létre az egyes állapotok között, és modellezi a variációkat az állapotokhoz rendelt jelek időtartama alatt. A balról–jobbra rejtett Markov modell esetében  $a_{ij} = 0$ , minden  $i > j$  esetén, ezért ez az  $\mathbf{A}$  mátrix egy felső háromszög mátrix.
- Állapot tapasztalati vektorok  $\{\mu_{i1}, \mu_{i2}, \dots, \mu_{iM}, i = 1, \dots, N\}$ . Minden állapotra egy halmaz  $M$  prototípus vektor, a jel terének centroidjainak modellezésére minden állapotban.
- Állapot tapasztalati vektor valószínűségi modellje. Lehet diszkrét az  $M$  prototípusvektorokból, azok hozzárendelt sűrűségfüggvényével  $\mathbf{P} = \{P_{ij}(\cdot); i=1, \dots, N, j=1, \dots, M\}$ , vagy folytonos (általában normális)  $\mathbf{F} = \{f_{ij}(\cdot); i=1, \dots, N, j=1, \dots, M\}$  eloszlásfüggvény modell.
- Kezdőállapot valószínűségi vektor  $\boldsymbol{\pi} = [\pi_1, \pi_2, \dots, \pi_N]$

2.2.2.3. A rejtett Markov modell betanítása

Az első lépés, hogy adatbázist alakítsunk ki a modell számára, nagy számú minta vételezésével a vizsgálat tárgyát képező jelből. Tegyük fel, hogy az adatbázis  $L$  vektor értékekkel bíró szekvenciából áll  $[\mathbf{X}]=[\mathbf{X}_k; k=0, \dots, L-1]$ , és minden szekvenciában  $\mathbf{X}_k=[x(t); t=0, \dots, T_k-1]$  változó számú  $T_k$  vektor van. A paraméterek megfelelő beállításához a modell statisztikájának adathalmazát kell előállítani. Valószínűségi értelemben a modell alkalmassága a modell utólagos valószínűségével  $P_{M|X}(M|\mathbf{X})$  adható meg,  $M$  modell  $\mathbf{X}$  adathalmazára. Ez a betanító folyamat az  $M$  modell utólagos valószínűségének és az  $[\mathbf{X}]$  adathalmaznak a maximalizálását célozza, a Bayes szabály alkalmazásával:  $P_{M|X}(M|\mathbf{X}) = \frac{1}{f_X(\mathbf{X})} f_{X|M}(\mathbf{X}|M)P_M(M)$ ,

ahol a nevezőben álló  $f_X(\mathbf{X})$  egy normalizáló hatással bír, és  $P_M(M)$  az  $M$  modell korábbi valószínűsége. Az iménti egyenlet maximalizálása egyenértékű a valószínűségi függvény maximalizálásával, mely a következő egyenlet segítségével fejezhető ki  $\mathbf{X}$  vektorszekvenciára és  $M$  modellre:

$$f_{X|M}(\mathbf{X}|M) = \sum_s f_{X|S,M}(\mathbf{X}|s, M)P_{s|M}(s|M),$$

ahol  $f_{X|S,M}(\mathbf{X}(t)|s(t), M)$  jelöli az eloszlásfüggvényt az  $M$  modell  $\mathbf{X}$  szekvenciájának  $\mathbf{s}=[s(0), s(1), \dots, s(T-1)]$  állapotában, mely a következőképpen fejezhető ki:

$$f_{X|S,M}(\mathbf{X}|s, M) = f_{X|S}(\mathbf{x}(0)|s(0))f_{X|S}(\mathbf{x}(1)|s(1))\cdots f_{X|S}(\mathbf{x}(T-1)|s(T-1)),$$

ahol az  $N$  állapot egyike lehet  $s(t)$   $t$  időpillanatban,  $f_{X|S}(\mathbf{x}(t)|s(t))$  pedig

$f_{X|S,M}(\mathbf{X}(t)|s(t), M)$  rövidített leírása az  $M$  modell  $\mathbf{x}(t)$  eloszlásfüggvényének  $s(t)$

állapotban. A  $\mathbf{s}$  állapot szekvencia Markov valószínűsége a következőképpen adható

$$P_{s|M}(s|M) = \pi_{s(0)}a_{s(0)s(1)}a_{s(1)s(2)}\cdots a_{s(T-2)s(T-1)}.$$

Az előző három egyenletből az utolsó kettőt behelyettesítve az elsőbe kapjuk a következőt:

$$\begin{aligned} f_{X|M}(\mathbf{X}|M) &= \sum_s f_{X|S,M}(\mathbf{X}|s, M)P_{s|M}(s|M) = \\ &= \sum_s \pi_{s(0)}f_{X|S}(\mathbf{x}(0)|s(0))a_{s(0)s(1)}f_{X|S}(\mathbf{x}(1)|s(1))\cdots a_{s(T-2)s(T-1)}f_{X|S}(\mathbf{x}(T-1)|s(T-1)) \end{aligned}$$

ahol az összegzés végighalad a összes  $\mathbf{s}$  állapot szekvenciát. A betanítási folyamatban az átmeneti valószínűségeket és a tapasztalati eloszlásfüggvények paramétereit az úgy kell kiszámítani, hogy iménti egyenletből a modell valószínűsége maximális legyen. A közvetlen maximalizálás a modell paramétereivel összefüggésben nem triviális megoldás. Következésképpen az egyenlet számításigénye egy megfigyelési



---

szekvencia  $T$  darab vektorára  $O(N^T)$  lesz. Ez még kis számok esetében is nagyon nagy számításigényt jelent. Az állapot–idő diagram ismétlődő struktúrái azt mutatják, hogy sok az ismétlődő számítás az egyenletben, melyet egy hatékony megvalósítással el tudunk kerülni.

### 3. Zajsűrési módszerek

#### 3.1. Zajzár

Érdekes, bár nem túl hatékony megoldás a jelszint folyamatos vizsgálata, és ha az elér egy kellően alacsony értéket, akkor a jelet teljesen kikapcsolva, csendet kapunk eredményül. Ezzel nem csak a zajt állítjuk meg, hanem az alacsony jelszintű hasznos jelet is, és a jel zajterhelése továbbra is megmarad, ha már egy adott szint felett van az értéke. Ez a módszer megoldást jelenthet olyan esetekben, mikor egy berendezés öngerjesztő folyamatának szeretnénk gátat szabni. Ilyen előfordulhat például mikrofonok esetén, melyek egy légtérben vannak az erősített jelüket megszólaltató hangsugárzókkal, és elhelyezéssel már nem tudunk a problémán segíteni (nagy hangerőre van szükség, és bárhová helyezzük a mikrofont, egy zárt térben, beindul az oszcilláció). Másik szemléletes példa az elektromos gitárok torzító berendezése, melynek alapvető működéséhez tartozik az átlagosnál nagyobb erősítés. Ilyenkor az érkező jel kikapcsolása a húrok öngerjesztő rezgésének megakadályozását szolgálhatja. Egy ilyen berendezés megfelelő beállítása nehéz, főleg, ha több vagy többféle hangforrás is van, ráadásul a valódi problémára nem jelent megoldást.

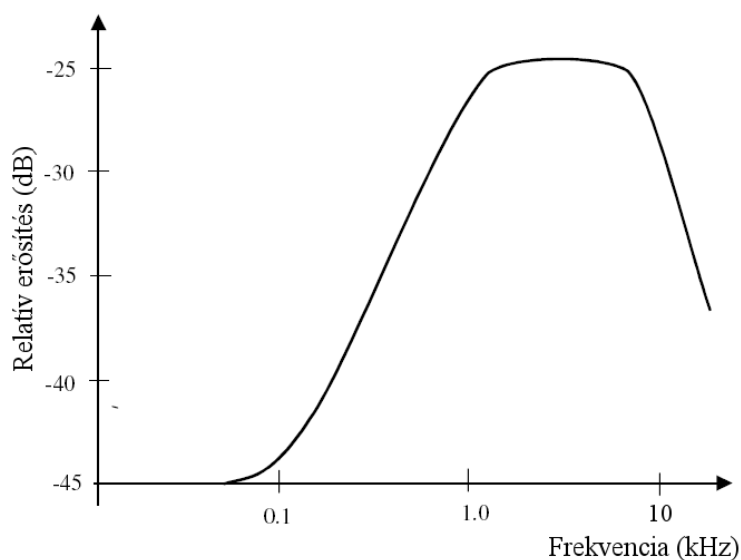
#### 3.2. Analóg valós idejű zajsűrők

##### 3.2.1. Dolby zajsűrő

A Dolby zajsűrő rendszerek úgy működnek, hogy emelik a magasabb frekvenciákon a hasznos jel amplitúdóját, ezzel javítva a jel–zaj arányt ezen frekvenciákon. A jel nagy része főként az alacsonyabb frekvenciákra – 2 kHz alá – koncentrálódik. A magasabb frekvenciájú jelek, melyek a minőséget és az élményt hordozzák viszonylag alacsonyabb energiájúak, és tovább degradálódnak a zajjal. Például egy mágneses szalagra rögzített jelsorozat esetében a szalag zaja és a korrekciós előerősítő fokozat zaja jelentős magasabb frekvenciájú „sistergést” állít elő. Lejátszásakor ezért ezeken a frekvenciákon romlik a jel–zaj arány – tehát nem lesz egyenletes a teljes átvitt sávban, vagyis az alacsonyabb frekvenciákon jobb, a magasabbakon rosszabb értéket kapunk. Tehát a magasabb frekvenciákon kevesebb hallható hasznos jelet kapunk. A Dolby zajsűrő tehát azon az alapelven működik, hogy a viszonylag kisebb energiájú magasabb frekvenciájú jeleket kiemeli, erősíti a hanganyag felvétele során, a többi jelet pedig változatlanul hagyja. Felvétel közben a jel

egy kiemelés és egy dinamika kompresszió keverékén megy keresztül, melyet lejátszáskor vissza kell alakítani az eredeti állapotába. A kódoló és a dekódoló rendszernek jól illesztettnek kell lennie, a torzítások elkerülése érdekében.

A Dolby laboratóriumban többféle ilyen szűrőt fejlesztettek ki, melyek a Dolby A, Dolby B, és Dolby C. Ezek alapjában véve a sávok számában és az alkalmazott kiemelő stratégiájában különböznek egymástól. A Dolby A a professzionális felhasználás számára lett kifejlesztve, és négy nagy részre osztja a frekvenciaspektrumot – az első sáv 0–80 Hz (alul áteresztő szűrő), a második sáv 80–3000 Hz (hullámsávszűrő), a harmadik sáv 3 kHz körüli felül áteresztő szűrő, a negyedik sáv pedig egy 9 kHz körüli felül áteresztő szűrő segítségével jött létre. A kódoló minden hullámsáv erősítését alkalmazkodóan igazította, az alacsony energiájú jelek kiemelése érdekében. A Dolby A rendszer gondoskodik a 10–15 dB-es erősítésről minden sávban, ha a jel 45 dB-nél nagyobb értékkel esik a maximális rögzítési jelszint alá. A kereskedelmi forgalomba kerülő készülékekben a Dolby B és a Dolby C zajsűrőket alkalmazták. Ezek a rendszerek az imént ismertetett Dolby A rendszerrel szemben már csak két hullámsávon dolgoznak, a négy helyett. A



17. ábra A Dolby C rendszer elő-kiemelő sávtávitele

Dolby B rendszer 10 dB-es emelést biztosít alacsony jelszint esetén (nagyobb, mint 45 dB eltérés a maximális rögzítési jelszinthez képest), a Dolby C rendszer pedig 20 dB-es emeléssel dolgozik, ahogy az a következő ábrán látható.

Ez a módszer csak analóg jelek esetében egy jól meghatározott zajforma csökkentésére alkalmas,

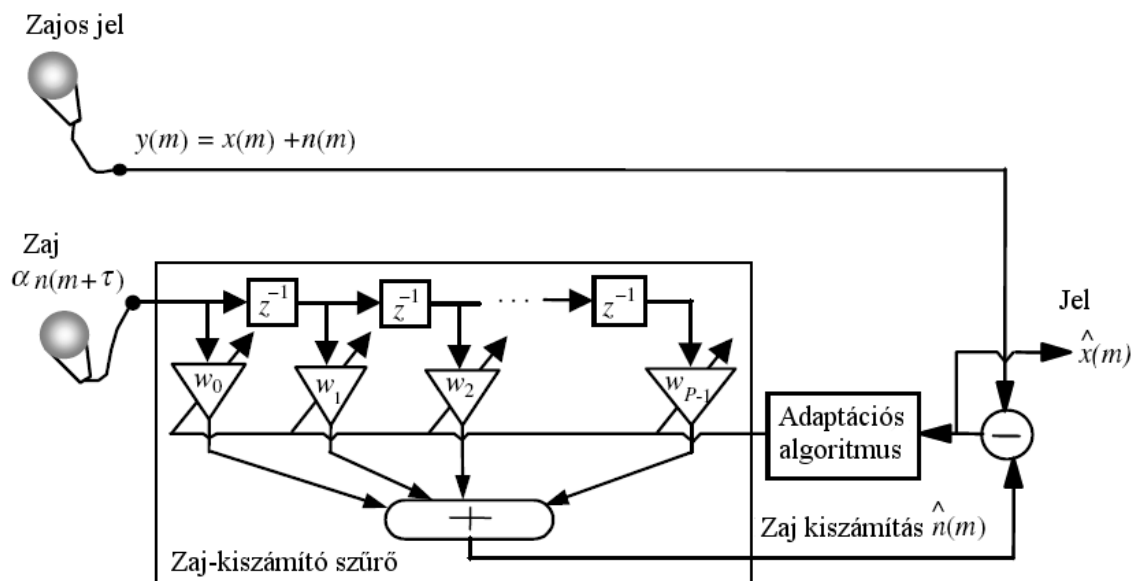
ezért ez nagyon behatárolja az alkalmazási területet. Ezért alkalmazzák ezt a módszert szalagos rögzítő eszközök – úgy mint kompakt kazettás magnetofonok – esetében. Az ilyen készülékek esetében jellemző, hogy a szalagon található mágneses részecskék rendezettsége és az előmágnesezés milyensége befolyásolja a hangminőséget, továbbá a magasabb frekvenciájú hangjelek – az alkalmazott mágne-

szehető anyag függvényében – kisebb jelszinten érkeznek vissza a szalagról, mely egy korrekciós előerősítő fokozat alkalmazását teszi szükségessé, ezáltal az elektronikus rendszernek is lesz egy zaja, melyek együttesének „szűrésére” alkalmazható a Dolby A, B és C rendszer.

### 3.2.2. Zajérvénytelenítés

Egy zajos környezetből (mint például egy mozgó autó) származó beszéd esetén a hasznos jelet a környezet zajával együtt észleljük. A mérőrendszerek esetében az információhordozó jel gyakran a berendezést körülvevő környezet zajával szennyezett. A zajjal együtt észlelt  $y(m)$  jel a következőképpen modellezhető:  $y(m) = x(m) + n(m)$ , ahol  $x(m)$  jelenti a hasznos jelet,  $n(m)$  jelenti a zajt,  $m$  pedig a diszkrét idő mérőszámot. Néhány esetben, például mikor mobiltelefont használunk egy mozgó autóban, vagy rádiókommunikációs eszközt egy repülőgép pilótafülkéjében, talán lehetséges, hogy kimérhető és kiszámítható a környezeti zaj pillanatnyi amplitúdója, egy irányított mikrofon használatával. Ezután a az  $x(m)$  hasznos jel visszaállítható lehet a zajos jel és a zaj különbségeként.

Az adaptív zaj-érvénytelenítő rendszer – két bemenettel rendelkezik, melyekből az egyik a környezeti zajt érzékeli, a másik pedig a zajjal terhelt hasznos jelet – alkalmazható beszédhang kiemelésre. Ebben a rendszerben az egyik irányított mikrofon a zajos  $x(m) + n(m)$  jelet érzékeli, míg a másik irányított mikrofon az előzőtől egy adott távolságra elhelyezve a zajt  $\alpha \cdot n(m + \tau)$  méri. Az a csillapítási szor-



18. ábra Adaptív zajérvénytelenítő elvi felépítése

zótényező és a  $\tau$  időkéleltetési érték meglehetősen túlegyszerűsített modellt ad a mikrofon elhelyezésétől eltérő helyről származó és eltérő zajok hatásainak modellezésére. A második mikrofonból érkező zajt egy adaptív digitális szűrővel dolgozzuk fel, annak érdekében, hogy amplitúdója megegyezzen az első mikrofonban érzékelt hang zajszennyezésének amplitúdójával, ezután pedig ezt a jelet kivonjuk zajos jelből, a zaj érvénytelenítése céljából. Ez a berendezés a zaj alacsonyabb frekvenciáin működik hatékonyan, de szenved a jelek nem állandó jellegétől, és a túlegyszerűsített feltevéstől mely szerint lineáris szűrővel modellezhető a zaj szóródása és terjedése a térben. A berendezés felépítését a 18. ábra szemlélteti.

Sok esetben, például egy telekommunikációs rendszerben alkalmazott vevőkészülék esetében nem ismerjük a zajszennyezés pillanatnyi értékét, csak a zajos jel adott. Ezen esetekben nem érvényteleníthető a zaj, csak mértéke csökkenthető, a jel és a zaj statisztikáját felhasználva. A Wiener szűrő alkalmas ilyen jellegű zajcsökkentési eljárásra, ahol csak a zajos jel adott.

### **3.3. Digitális jelek feldolgoása**

#### *3.3.1. Átalakítások*

##### 3.3.1.1. Analóg–digitális átalakítás

Az analóg jelfeldolgozás során valamilyen folytonosan változó fizikai mennyiséget alakítunk elektronikusan feldolgozható jellé. Ilyen átalakítás történik például egy mikrofonban, ahol a levegő rezgéseit, nyomásváltozásait alakítjuk elektromos feszültséggé, melyet a későbbiekben elektronikus jelerősítő berendezésekkel tudunk tovább feldolgozni. Ezek a jelek még elég sérülékenyek, könnyen terhelődhetnek zajjal, és ezen zajok eltávolítása az analóg jelekből nagyon nehéz feladat. Ha a jelet szeretnénk megszabadítani az additív zajoktól, akkor az első lépés, amit meg kell tennünk a jel átalakítása digitális jellé. A digitális jel fontos jellemzője, hogy egy előre meghatározott értékhalmból vehet fel csak értékeket. Ezen jelek esetében a feldolgozás már matematikai modellek alapján digitális számítógépek segítségével történhet. Továbbá nagyon előnyös tulajdonságaik közé tartozik, hogy sokkal kevésbé érzékenyek a külső additív zajokra, a diszkrét értékek miatt. A legegyszerűbb ilyen digitális jelek kétértékűek, vagy más néven bináris jelek. A mai korszerű kommunikációs rendszerekben gyakran alkalmaznak olyan készülékeket, melyek az analóg jelet először csak digitálissá alakítják, majd ez a digitális jelet

alakítják tovább egy bitfolyammá, melynek értelmezése a vevő oldalon történik. Ez a bit soros adatátvitel, napjaink egyik legkedveltebb átviteli formája.

Az analóg–digitális átalakítás vagy konverzió (Analog to Digital Conversion – ADC) úgy történik, hogy egy analóg feszültségjelet alapul véve abból adott időközönként mintát veszünk. Ez a digitális jel egyik fontos jellemzője, a *mintavételezési frekvencia*. A legjobb minőségű digitális jelet akkor kapnánk, ha egy időegység alatt végtelen sokszor vennénk mintát az analóg jelből, de ez technikailag kivitelezhetetlen és értelmetlen is. A mintavételezés a gyakorlatban rövid időközönként történik:

$T_s = \frac{1}{F_s}$ , ahol  $T_s$  az időközt,  $F_s$  pedig a frekvenciát jelöli. Ez a *mintavételezési eljárás*

az  $x(t)$  analóg jelet egy  $x(m)$  impulzussorozattá alakítja.

Az átalakítás következő lépése a kvantálás, amikor a mintavételezés után kapott impulzusokat egy diszkrét értékhalmozra képezzük. A halmaz elemeit egy analóg jeltartomány meghatározott számú részekre osztásával kapjuk meg. A kvantálás során a mintavételezés eredményeképpen kapott impulzussorozat eleméről eldöntjük, hogy melyik két érték közé (melyik tartományba) esik, és előállítjuk a  $x(m)$  jelből az  $y(m)$  jelet, mely a kvantálás eredményeképpen keletkezik. Ez a digitális jel másik fő jellemzője, amit *kvantálási hossz*nak mondunk. A jel pontossága és a visszaalakításkor kapott eredmény nagymértékben függ a tartományok számától, így a kvantálási hosszról.

A következő lépés a kódolási művelet. Ebben a műveletben minden mintavételi pontban felvett jelhez a kvantálás végeredményének megfelelő bináris kódszót rendelünk hozzá. Ezzel az eljárással minden  $y(m)$  jelhez bináris kódszót rendeltünk, tehát végeredményként kaptunk egy digitálisan feldolgozható jelet.

### 3.3.1.2. Digitális–analóg átalakítás

A digitális jel analóg jellé történő alakításakor (Digital to Analog Conversion – DAC) először a vett bináris jelsorozatot kell helyreállítani, amennyiben az szükséges. A jelsorozat visszaállítását követően a kódszavakból előállítható az  $y'(m)$  függvény, melyet egy alul-áteresztő szűrővel szűrve megkapható az eredeti  $x(t)$  jelet közelítő  $x'(t)$  analóg jel.

A közelítés javítására alkalmazhatunk interpolációt, mely a két függvény hasonlóságát növeli és csökkenti a torzítást. Az analóg jel mintavételezett formájából történő visszaállítás feltétele a Nyquist elméletben megadott. A Nyquist elmélet azt mondja ki, hogy egy sávkorlátozott jel esetében annak  $F_c$  legnagyobb frekvenci-

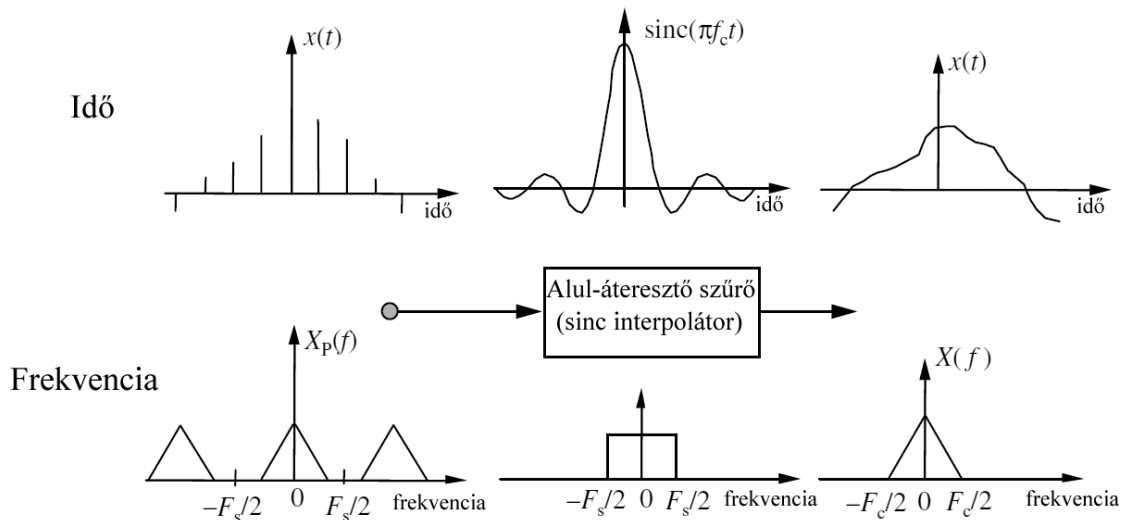
ájú összetevője is visszaállítható abban az esetben, ha a mintavételezés másodpercenkénti frekvenciája nagyobb volt a frekvencia kétszeresénél, vagyis  $F_s > 2F_c$ . Tegyük fel, hogy egy sávkorlátozott  $x(t)$  folytonos jel  $F_s$  másodpercenkénti mintavételezési frekvenciával mintavételezett. A diszkrét időközönként mintavételezett jel  $x(m)$  a következő szorzattal fejezhető ki:  $x(m) = x(t)p(t) = \sum_{m=-\infty}^{\infty} x(t)\delta(t - mT_s)$ , ahol

$p(t) = \delta(t - mT_s)$  a mintavételező függvény, és  $T_s = \frac{1}{F_s}$  a mintavételezési idő intervallum hossza. Vegyük az iménti egyenletet Fourier transzformálva, mely megmutatja, hogy a mintavételezett jel spektruma a következő egyenlettel írható fel:

$X_s(f) = X(f) * P(f) = \sum_{k=-\infty}^{\infty} X(f + kF_s)$ , ahol  $X(f)$  az  $x(t)$  jel spektruma, és  $P(f)$  a  $p(t)$

mintavételező függvénye, a  $*$  pedig a konvolúciós szorzás műveletét jelöli. Az iménti egyenlet állítása szerint a mintavételezett jel spektruma az alapsáv  $X(f)$  spektrumából és annak egyenletes  $T_s = \frac{1}{F_s}$  közönkénti ismétléséből vagy leképezéséből

áll össze. Ha a mintavételezési frekvencia felette van a Nyquist aránynak, akkor az alapsáv  $X(f)$  spektruma nincs átfedésben a  $X(f \pm kF_s)$  képeivel, és az eredeti jel



19. ábra Folytonos jel visszaállítása mintáiból. A frekvenciatartomány interpolációja az alul-áteresztő szűrővel ekvivalens

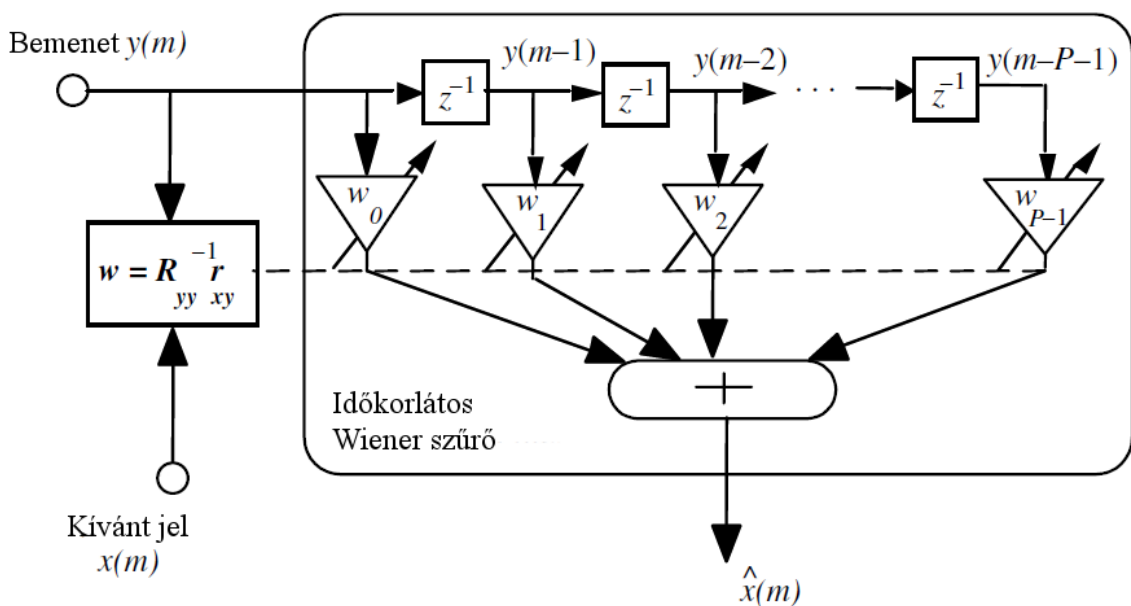
A gyakorlatban az  $F_s$  mintavételezési frekvencia nagyobb mint  $2F_s$ , általában  $2,5F_s$ , annak érdekében, hogy az átmeneti sáv szélességet hozzáigazítsuk az interpoláció alul-áteresztő szűrőjéhez.

### 3.3.2. A Wiener szűrő

A Wiener elméletet kidolgozójáról, Norbert Wienerről nevezték el, aki megfogalmazta az adatfüggő lineáris legkisebb négyzetek módszerére épülő szűrő elméletét. A Wiener szűrő központi szerepet játszik alkalmazásainak széles körében, mint például a visszhang érvénytelenítés, jelvisszaállítás, csatorna kompenzáció és rendszerazonosítás. A szűrő együtthatóit úgy kell kiszámítani, hogy a szűrő kimenő jele a lehető legjobban hasonlítson – vagyis különbségük négyzetének várható értéke legyen minimális – egy ismert úgynevezett megkívánt jelre. A Wiener elmélet feltételezi, hogy a jelek állandóak. Azonban ha a szűrő együtthatóit időnként újra kiszámítjuk, minden  $N$  elemű jelblokk után, a szűrő alakítja önmagát, a blokkokon belüli átlagos jellemzőkhöz, így blokk-alkalmazkodóvá válik. A blokk-alkalmazkodó (vagy szegmens-alkalmazkodó) szűrő alkalmazható olyan jelek esetében, mint például az emberi beszéd.

#### 3.3.2.1. A legkisebb négyzetek becslése

Wiener eredeti munkájában a folyamatos idejű legkisebb négyzetek módszerét adta meg, interpolációval. A módszert később kiterjesztette diszkrét időre is, mely egyszerűen megoldható, és a gyakorlatban jobban hasznosítható a digitális jelpro-



20. ábra Wiener szűrő felépítése



cesszorokban. A szűrő lehet végtelen hosszú jel feldolgozását célzó, vagy véges jel-hosszúságot figyelembevevő. Általában a végtelen szűrőnek egy halmaz nem lineáris egyenlet az eredménye, a véges hosszúságú jelre alkalmazott szűrőnek pedig egy halmaz lineáris egyenlet, mely egy zárt alakú megoldás. A véges szűrő viszonylag egyszerűbb számítást tesz lehetővé, működéséből adódóan stabilabb és praktikusabb is. Hátránya a végtelen szűrővel szemben, hogy a kívánt eredmény elérése érdekében nagyon sok együttható szükséges.

A 20. ábrán szemléltetett Wiener szűrő egy  $\mathbf{w}$  együttható vektorral megvalósított. A szűrő kap egy  $y(m)$  bemeneti jelet, és készít egy  $\hat{x}(m)$  kimeneti jelet, mely a legkisebb négyzetes eltérést mutatja az  $x(m)$  megkívánt vagy céljelhez képest. A szűrő bemenete és kimenete közötti összefüggés a következő egyenlettel írható fel:

$$\hat{x}(m) = \sum_{k=0}^{P-1} w_k y(m-k) = \mathbf{w}^T \mathbf{y}, \text{ ahol } m \text{ a diszkrét idő mérőszám, a szűrő bemeneti}$$

jele:  $\mathbf{y}^T = [y(m), y(m-1), \dots, y(m-P+1)]$ , és a paramétervektor (a Wiener szűrő együttható vektora)  $\mathbf{w}^T = [w_0, w_1, \dots, w_{P-1}]$ . A szűrési művelet kétféle egymással ekvivalens alakban fejthető ki, konvolúciós összegként, és a belső vektorok szorzataként. A hibajel  $e(m)$  a kimenőjel  $\hat{x}(m)$  és a megkívánt jel  $x(m)$  különbségeként számítható ki a következő összefüggéssel:  $e(m) = x(m) - \hat{x}(m) = x(m) - \mathbf{w}^T \mathbf{y}$ . Ebből az egyenletből kiderül, hogy egy adott  $y(m)$  bemenőjelre és  $x(m)$  megkívánt jelre a hibajel a  $\mathbf{w}$  együtthatóvektortól függ. A hibajel és az együtthatóvektor közötti összefüggés kiderítése érdekében kifejtve az előző egyenletet a következőt kapjuk:

$$\begin{pmatrix} e(0) \\ e(1) \\ e(2) \\ \vdots \\ e(N-1) \end{pmatrix} = \begin{pmatrix} x(0) \\ x(1) \\ x(2) \\ \vdots \\ x(N-1) \end{pmatrix} - \begin{pmatrix} y(0) & y(-1) & y(-2) & \dots & y(1-P) \\ y(1) & y(0) & y(-1) & \dots & y(2-P) \\ y(2) & y(1) & y(0) & \dots & y(3-P) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y(N-1) & y(N-2) & y(N-3) & \dots & y(N-P) \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_{P-1} \end{pmatrix}.$$

Ez az egyenlet  $N$  mintára,  $y(m)$  és  $x(m)$  jelekre lett felírva. Tömörített alakban a következőképpen írható fel:  $\mathbf{e} = \mathbf{x} - \mathbf{Y}\mathbf{w}$ , ahol  $\mathbf{e}$  a hibavektor,  $\mathbf{x}$  a kívánt jel vektor,  $\mathbf{Y}$  a bementi jel mátrix, és  $\mathbf{Y}\mathbf{w}$  szorzat pedig a  $\hat{x}$  kimenőjel vektor. Feltételezzük, hogy  $P$  kiinduló minta ismert, vagy hanem akkor 0 értékűnek tekintendő. Ha az előbb felírt mátrixegyenletben az  $N$  és a  $P$  értéke egyenlő, akkor egy négyzetes mátrixegyenletet kapunk, és ez az egyetlen megoldás, ahol a számítási hiba értéke nulla. Ha  $N < P$  a minták száma  $N$  elégtelen mennyiségű ahhoz, hogy elérjük az egyedülálló megoldást a szűrő együtthatóinál, ezért ebben az esetben végtelen sok megoldás létezik nulla számítási hibával és a mátrixegyenletet határozatlannak mond-

juk. A gyakorlatban a jelminták száma sokkal nagyobb, mint a szűrő hossza – vagyis  $N > P$  –, ebben az esetben a mátrixegyenletet túlhatározottnak mondjuk, és létezik egy egyedi megoldás, általában nulla számítási hibával. Mikor  $N > P$  a szűrő együtthatói úgy lesznek kiszámítva, hogy az átlagos hiba minimális legyen, mint például a hiba átlagos értékének abszolút értéke  $E[|e(m)|]$ , vagy a hiba négyzetének középértéke  $E[e^2(m)]$ , ahol  $E[.]$  a várható érték operátor. A hibafüggvény kiválasztása hatással van a megoldás optimális mivoltára, és a számítási komplexitására. A Wiener elméletben az objektív kritérium a legkisebb négyzetes hiba (LSE<sup>2</sup>) a szűrő kimeneti jele és a kívánt jel között. A legkisebb négyzetes hiba kritérium optimális a normális eloszlást követő jelek esetében. A korlátozott idő intervallumra alkalmazott legkisebb négyzetes hiba kritériumú Wiener szűrő egy lineáris, zárt formulához vezet. A Wiener szűrő együtthatói egy átlagos négyzetes hiba függvény  $E[e^2(m)]$  minimalizálásával eszközli ki, tekintettel a szűrő  $\mathbf{w}$  együttható vektorára. A számítási hiba négyzetének középértéke a következő egyenlettel adható meg:  $E[e^2(m)] = E[(x(m) - \mathbf{w}^T \mathbf{y})^2] = E[x^2(m)] - 2\mathbf{w}^T E[\mathbf{y}x(m)] + \mathbf{w}^T E[\mathbf{y}\mathbf{y}^T] \mathbf{w} = r_{xx}(0) - 2\mathbf{w}^T \mathbf{r}_{yx} + \mathbf{w}^T \mathbf{R}_{yy} \mathbf{w}$ , ahol  $\mathbf{R}_{yy} = E[\mathbf{y}(m)\mathbf{y}^T(m)]$  a bemenő jel autokorrelációs<sup>3</sup> mátrixa,  $\mathbf{r}_{xy} = E[x(m)\mathbf{y}(m)]$  pedig a bemeneti jel és a kívánt jel kereszt-kölcsönösségi vektora. Az előző kifejezést kifejtve a következő egyenletet kapjuk:

$$E[e^2(m)] = r_{xx}(0) - 2 \sum_{k=0}^{P-1} w_k r_{yx}(k) + \sum_{k=0}^{P-1} w_k \sum_{j=0}^{P-1} w_j r_{yy}(k-j), \text{ ahol } r_{yy}(k) \text{ és } r_{yx}(k) \text{ az } \mathbf{R}_{yy}$$

autokorrelációs mátrix és az  $\mathbf{r}_{xy}$  kereszt-kölcsönösségi vektor elemei külön-külön. Az előző egyenletből a korlátozott idő intervallumra alkalmazott szűrő négyzetes hibájának középérték számítására egy másodfokú függvény alkalmazható a  $\mathbf{w}$  együttható vektor értékeire, mely függvénynek egyetlen minimuma van. Például ha szűrőnek csak két együtthatója van ( $w_1, w_2$ ), akkor a négyzetes hiba középérték függvénye egy medence alakú képet mutat, egyetlen minimummal. A legkisebb négyzetes hiba jelenti az energiaminimumot.

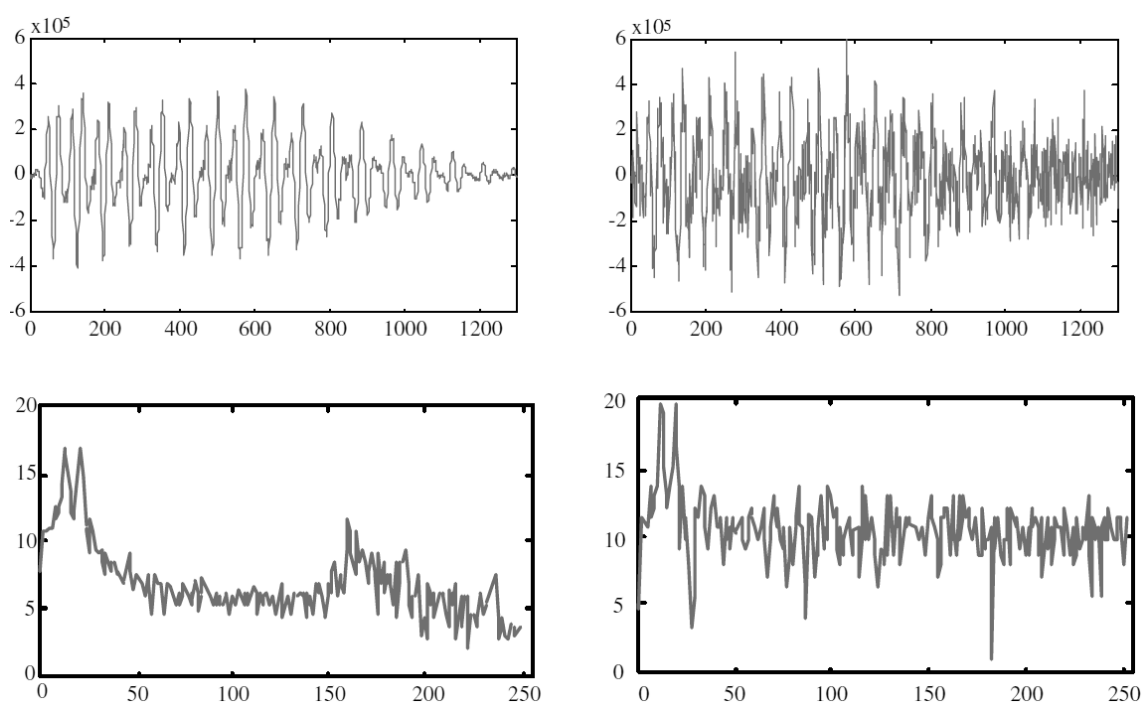
### 3.3.3. Zajcsökkentés spektrális kivonással

A spektrális kivonás módszerének lényege, hogy a zajos jelből veszünk egy olyan szakaszt, ahol a hasznos jel nincs jelen, tehát csak zaj észlelhető, és ezt a jelet egy megfelelő módszerrel kivonjuk a zajos jelből. Ez megtehető oly módon, hogy a jelek statisztikáit és spektrális paramétereit vesszük alapul. Ezzel a módszerrel adott esetben jelvisszaállítás is lehetséges.

A problémára a legegyszerűbb megoldást jelentené, ha külön rendelkezésre állna a zajos jel és a zaj. Ebben az esetben egy egyszerű művelettel eltávolítható lenne a zaj. A zajérvénytelenítő, melynek működését már korábban kifejtettem pontosan ilyen elven működik. Általában azonban csak a zajos jel áll rendelkezésre, tehát nincs lehetőség a zaj érvénytelenítésére, de előfordulhat, hogy a zaj hatását csökkenteni tudjuk a hasznos jel spektrumán.

### 3.3.3.1. A spektrális kivonás módszere

A 21. ábra olyan jeleket mutat, melyek zajjal terhelték, de nincs lehetőség a zaj érvénytelenítésére. A hasznos jel spektrumának középvértéke növelhető úgy, hogy a

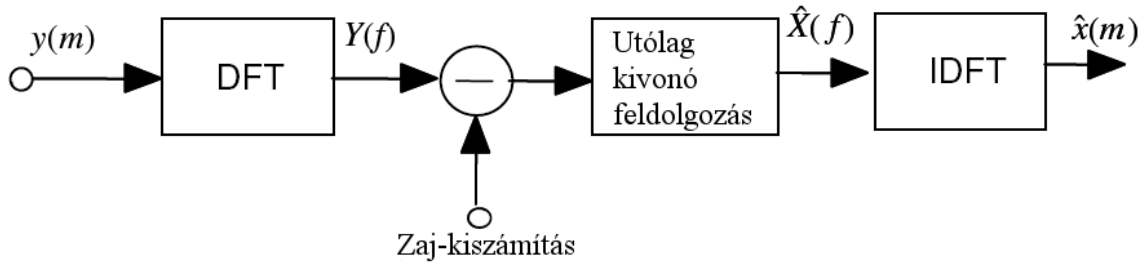


21. ábra Zajjal terhelt jelek alakjai az idő függvényében

zajos jel spektrumának középvértékéből kivonjuk a zaj spektrumának középvértékét. A zajos jel időben a következőképpen modellezhető:  $y(m)=x(m)+n(m)$ , ahol  $y(m)$  jelöli a zajjal terhelt jelet,  $x(m)$  a hasznos jelet,  $n(m)$  a zajt,  $m$  pedig a diszkrét idő mérőszám. Egy frekvenciatartományban az előző egyenlet a következőképpen fejthető ki:  $Y(f)=X(f)+N(f)$ , ahol  $Y(f)$  az  $y(m)$  zajos jel Fourier transzformáltja,  $X(f)$  az  $x(m)$  hasznos jel Fourier transzformáltja,  $N(f)$  az  $n(m)$  zaj Fourier transzformáltja, és  $f$  a frekvenciaváltozó. A spektrális kivonásban az  $y(m)$  zajos jelet vesszük, ezt egy átmeneti tárolóba helyezzük, majd felosztjuk  $N$  egyenlő részre. Minden szegmens esetében Hann vagy Hamming ablak segítségével szűrjük a spektrumot, ezután transzformáljuk diszkrét Fourier transzformációval (DFT) az összes  $N$  darab spektrális mintát. Az ablakok minden szegmens végén csökkentik annak hatását,

hogy a vett jelet feldaraboltuk, vagyis nem folytonos jeleket illesztünk össze. Az ablakozott jel így adható meg:  $y_w(m) = w(m)y(m) = w(m)[x(m) + n(m)] = x_w(m) + n_w(m)$ . Az alakozási eljárás a frekvenciatartományokra az alábbi módon fejthető ki (DFT segítségével):  $Y_w(f) = W(f) * Y(f) = X_w(f) + N_w(f)$ , ahol a  $*$  operátor a konvolúciós szorzást jelenti.

A következő ábra a spektrális kivonás módszerének blokkdiagramját szemlélteti. (Az IDFT jelölés a diszkrét Fourier transzformáció inverzét jelenti)



22. ábra A spektrális kivonás egyszerűsített blokkdiagramja

A spektrális kivonás egyenlettel a következő módon írható le:  $|\hat{X}(f)|^b = |Y(f)|^b - \alpha \overline{|N(f)|^b}$ , ahol  $|\hat{X}(f)|^b$  az eredeti jel spektrumának kiszámítása,  $|X(f)|^b$  és  $|N(f)|^b$  pedig az időátlagolt zaj spektrum. A spektrális kivonás nagyságához a  $b$  kitevő értékét 1-re, az erősségéhez pedig 2-re kell választani. Az  $\alpha$  paraméter a kivont zaj mennyiségét szabályozza. A teljes kivonáshoz  $\alpha=1$  értéket kell választani, a túlvonáshoz pedig  $\alpha>1$  érték lesz a megfelelő. Az időátlagolt zajspektrum beszerezhető a jelmentes szakaszokból, ahol csak kizárólag a zaj van jelen, nincs hasznos jel, és az alábbi összefüggés segítségével formalizálható:

$$\overline{|N(f)|^b} = \frac{1}{K} \sum_{i=0}^{K-1} |N_i(f)|^b, \text{ ahol } |N_i(f)|^b \text{ a spektrum } i\text{-edik zajkerete, és feltesszük,}$$

hogy a zaj  $K$  keretből áll, és  $K$  változó. Más esetben az átlagolt zajspektrum beszerezhető az elsőrendű digitális alul-áteresztő szűrő kimeneteként a következőképpen:

$$\overline{|N_i(f)|^b} = \rho \overline{|N_{i-1}(f)|^b} + (1 - \rho) |N_i(f)|^b, \text{ ahol az alul-áteresztő szűrő együtthatója}$$

$\rho$ , melynek tipikus értéke 0,85 és 0,99 közé szokott esni. Az időtartomány alapú jelvisszaállításhoz a nagyságspektrum  $|\hat{X}(f)|$  kiszámítását kombináljuk a zajos jel fázisával, majd ezt követően transzformáljuk időtartományba az inverz diszkrét Fourier transzformáció segítségével, mely a következőképpen írható fel:

$$\hat{x}(m) = \sum_{k=0}^{N-1} |\hat{X}(k)| e^{j\theta_Y(k)} e^{-j\frac{2\pi}{N}km}, \text{ ahol } \theta_Y(k) \text{ a zajos jel fázisa } Y(k) \text{ frekvencián. Ez a jel-}$$

viSSzaállító egyenlet feltételezi, hogy a hallható tartományba eső zaj főként a nagyságspektrum torzítása és a fázistorzítás miatt nagyrészt nem hallható. A szimulált fázistorzítás megfigyelhető hatásának becslése igazolta a feltevést. A zaj spektrumának változatossága következtében előfordulhat, hogy nagyság és az energiaspektrum szempontjából negatív eredményt kapunk a spektrális kivonással. Ennek végeredményeképpen nagyon valószínű, hogy csökkenni fog a jel-zaj arány. A negatív eredmények elkerülése érdekében a spektrális kivonás kimeneti jelét egy utófeldolgozó térképező függvénnyel feldolgozzuk, melyet a következőképpen formalizálhatunk:  $T[|\hat{X}(f)|] = \begin{cases} |\hat{X}(f)| & \text{ha } |\hat{X}(f)| > \beta |Y(f)| \\ \text{fn}[|Y(f)|] & \text{egyébként} \end{cases}$ . Például választhatunk egy

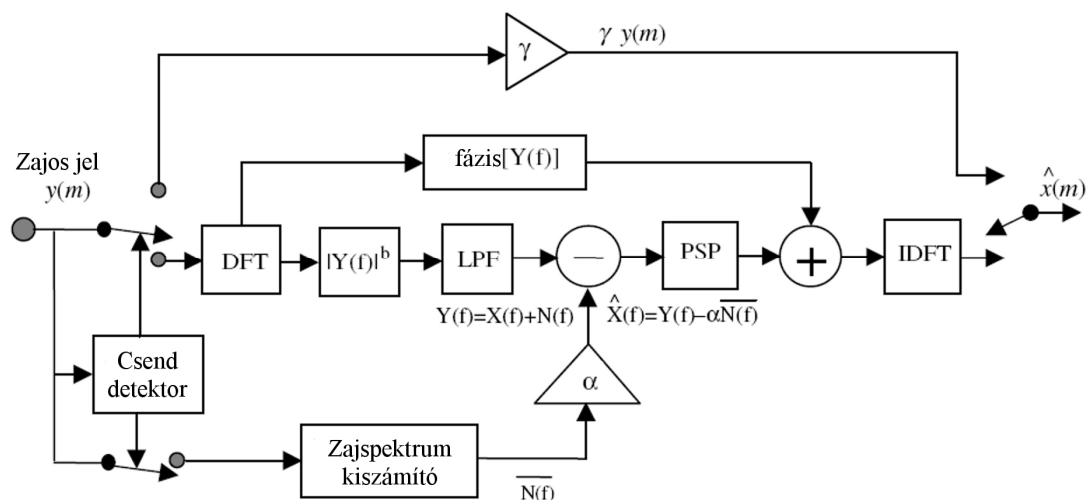
olyan szabályt, mely ha a  $|\hat{X}(f)| > 0,01 |Y(f)|$  (a nagyság spektrum szorzása 0,01-dal ekvivalens a  $-40\text{dB}$ -lel) érték igaz, akkor az  $|\hat{X}(f)|$  érték valamilyen függvénye legyen a zajos jelnek  $\text{fn}[|Y(f)|]$ . Ennek legegyszerűbb formája, hogy beállítunk egy  $\text{fn}[|Y(f)|] = \text{zajszint}$  értéket, ahol a zajszint pozitív konstans. Egy másik lehetőség, ha  $\text{fn}[|Y(f)|] = \beta |Y(f)|$ . Ez esetben:

$$T[|\hat{X}(f)|] = \begin{cases} |\hat{X}(f)| & \text{ha } |\hat{X}(f)| > \beta |Y(f)| \\ \beta |Y(f)| & \text{egyébként} \end{cases}$$

A spektrális kivonás energiaspektrum vagy nagyságspektrum tartományok alapján is implementálható. A két módszer hasonló, de elméletileg eredményeznek némi különbséget a várható teljesítményben.

### 3.3.3.2. A spektrális kivonó megvalósítása

Egy spektrális zajkivonó rendszer blokkdiagramját mutatja az alábbi ábra,



23. ábra Spektrális kivonó megvalósításának blokkdiagramja

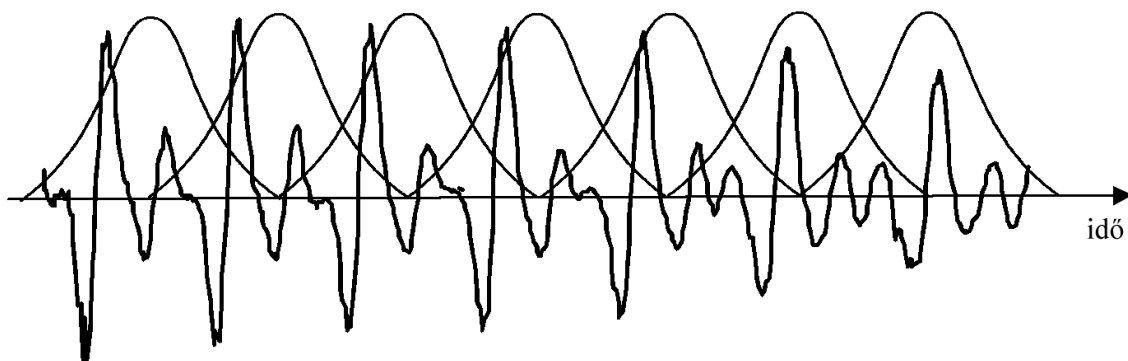
mely a következő alrendszeréből áll:

- a.) egy csend detektor a jelmentes időszakok megtalálására, a zaj spektruma ezen időtartamok alatt frissül.
- b.) a diszkrét Fourier transzformációt megvalósító rész (DFT) az időtartományt frekvenciatartományra alakítja, a DFT egységet egy nagyság operátor követi.
- c.) alul-áteresztő szűrő (Lowpass Filter, LPF), a zaj egyenetlenségeinek csökkentésére, melynek célja a zaj ingadozásaiból adódó feldolgozási torzítás csökkentése.
- d.) utófeldolgozó, a spektrális kivonás eredményeképpen keletkezett torzítások eltávolítására.
- e.) inverz diszkrét Fourier transzformációt végző rész, mely frekvenciatartományból viszi a feldolgozott jelet időtartományba.
- f.) csillapító berendezés  $\gamma$ , mely a zaj csillapítását végzi a csöndes részekben.

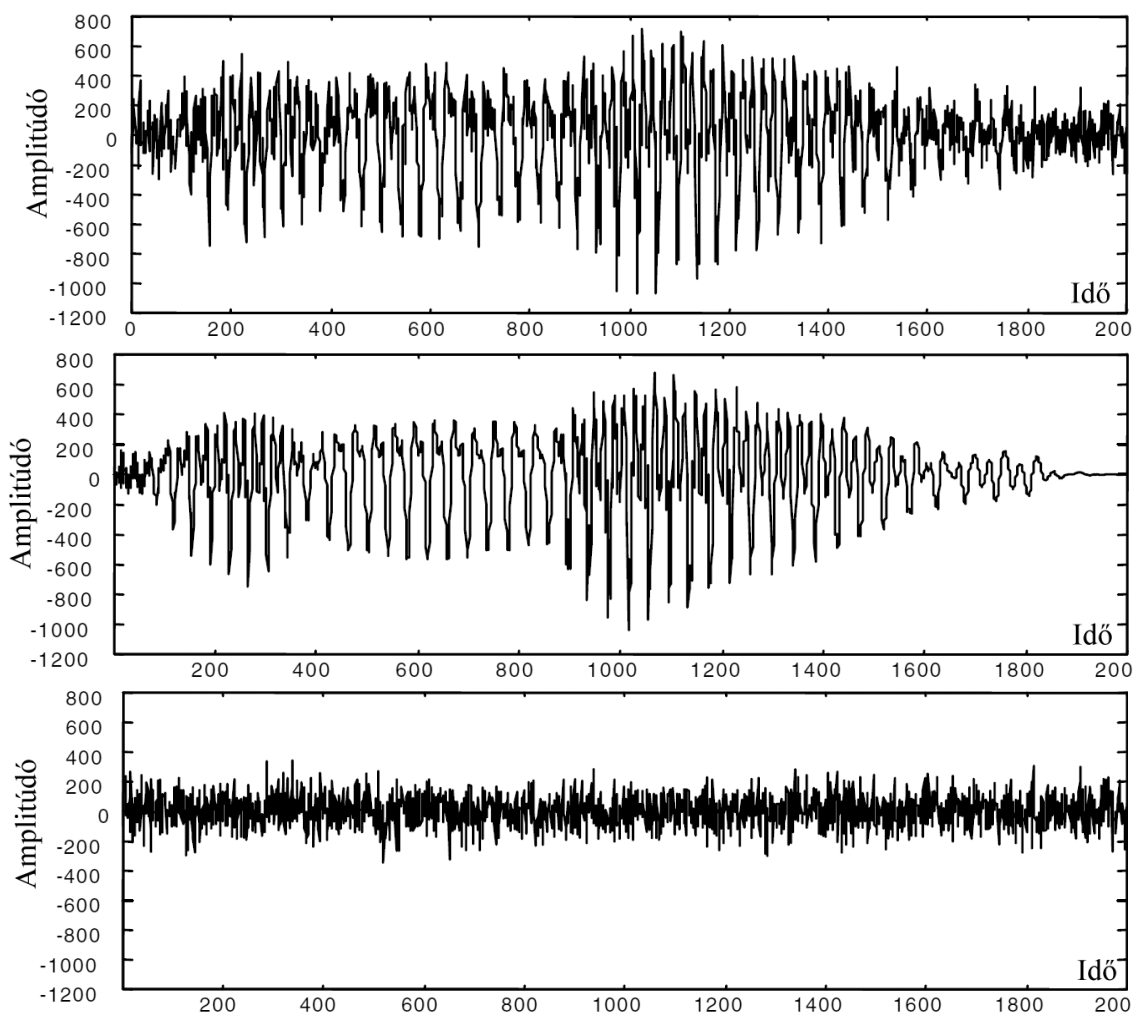
A diszkrét Fourier transzformáción alapuló spektrális kivonó blokkdiagramja egy algoritmus. A beérkező audiojelet átmeneti tárolóba rakjuk, majd  $N$  darab egymást átfedő mintablokkra osztjuk. Minden blokk Hann vagy Hamming ablakolt, és azt követően diszkrét Fourier transzformációval átalakítjuk frekvenciatartományra. A spektrális kivonás után a jel nagyságspektrumát kombináljuk a zajos jel fázisával, és visszaalakítjuk időtartományra. Minden blokkot egymáshoz illesztünk, úgy, hogy legyen átfedés a megelőző és a rákövetkező blokkal, és ez adja a végleges kimeneti jelalakot.

A blokk hosszának megválasztása a spektrális elemzéshez egy kompromiszsum, mely két ellentmondó szükséglet, az időfelbontás és a spektrális felbontás igényeit kell, hogy kielégítse. Tipikusan 5–50 ms közé esik a használt blokkok hossza. Ez egy 20 kHz-es mintavételezési frekvencián azt jelenti, hogy  $N$  értéke 100–1000 tartományba esik. A spektrum frekvenciafelbontása egyenesen arányos a minták számával ( $N$ ). Nagyobb  $N$  esetén a spektrum értékelése jobb lesz. Ez csak részben igaz a frekvenciatartomány alsó részére, mivel az alacsonyabb frekvenciájú részek az idő múlásával lassan változnak, és nagyobb ablak szükséges a stabil értékeléshez. Az ellentmondó szükséglet az audiojelek nem stacionárius természetéből adódnak, tehát az ablak méretének azért nem szabad túl nagyoknak lennie, hogy a rövid időtartamú események ne mosódjanak el.

Az ablakok átfedéseinek (24. ábra) fő feladata, hogy csillapítsák a blokkok végpontjain keletkező folytonossághiányokat, melyek a darabolásból adódnak. Bár

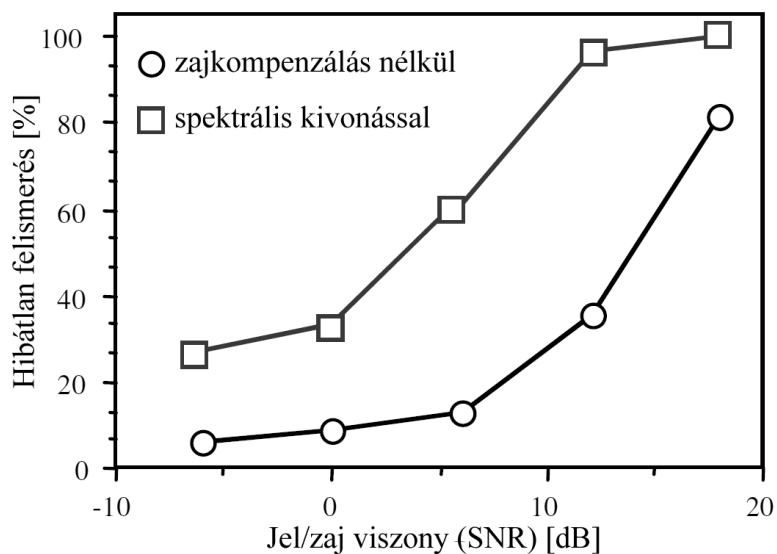


24. ábra A spektrális kivonásban alkalmazott ablakolási és átfedési folyamat létezik számos felhasználható ablaktípus, különböző karakterisztikákkal, mégis a legtöbb megvalósításban Hamming ablakokat használnak. A torzítás eltávolítás során – melyet a spektrális kivonás eredményezett – az utófeldolgozó algoritmus olyan információkat használ fel, mint a frekvenciacsatornák összefüggése az egymást követő blokkokban, valamint az jel eseményeinek és torzításának időtartamai.



25. ábra Zajos jel (felső), zajmentes jel (középső), és zaj (alsó)

A jel spektrális összetevőinek összefüggése – az időtengely mentén – részben szabályozható az ablakhossz megválasztásával és az átfedéssel. Ez az összefüggés növekszik az ablak hosszának csökkentésével és az átfedés növelésével. Az átfedés növelése azonban növeli a zajfrekvenciák összefüggését is, az időtengely mentén.



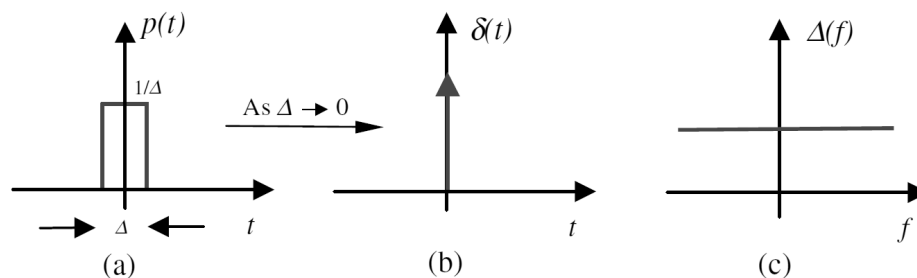
26. ábra A spektrális kivonás jelentősége a beszéd felismerésben

### 3.3.4. Az impulzív zaj eltávolítása

Az impulzív zajok gyors be/ki kapcsolásokhoz hasonló impulzusokat produkálnak és nagyon kiszámíthatatlan a megjelenésük, tehát bármely időpillanatban előfordulhatnak, és bármekkora időközrel. Ezen zajok eltávolítására és csökkentésére az előző részben kifejtett spektrális kivonás módszere nem alkalmas, mert itt a zajszint nem határozható meg.

#### 3.3.4.1. Impulzív zaj matematikai modellje

Az analóg impulzív zajt matematikai értelemben az alábbi ábra segítségével szemléltethetjük. Ha megvizsgáljuk az ábra (a) részén szemléltetett egységnyi te-



27. ábra (a) Egységnyi területű lökés, (b) impulzus, (c) az impulzus spektruma



rületű lökeshullám  $p(t)$  (négyzög jel) viselkedését, ahogy a lökés szélessége  $\Delta$  nullához tart, a lökés impulzussá válik. Az impulzus függvényt az ábra (b) része mutatja, és egy végtelenül kicsi időszélességgel rendelkező lökést a következőképpen

definiálhatunk:  $\delta(t) = \lim_{\Delta \rightarrow 0} p(t) = \begin{cases} \frac{1}{\Delta}, & |t| \leq \frac{\Delta}{2} \\ 0, & |t| > \frac{\Delta}{2} \end{cases}$ . Az impulzusfüggvény integrálja a kö-

vetkező:  $\int_{-\infty}^{\infty} \delta(t) dt = \Delta \times \frac{1}{\Delta} = 1$ . Az impulzusfüggvény Fourier transzformáltja felállítha-

tó:  $\Delta(f) = \int \delta(t) e^{-j2\pi ft} dt = e^0 = 1$ , ahol  $f$  a frekvenciaváltozó. Az impulzusfüggvényt

teszt jelleggel is szokták alkalmazni, a rendszer impulzusreakciójának megállapítására. Az ábra (c) részén mutatottak szerint az impulzus egy spektrálisan nagyon gazdag jel, minden frekvenciát azonos mennyiségben tartalmaz, tehát ezért alkalmas a tesztelésre.

Egy digitális impulzus  $\delta(m)$  (8. oldal 4. ábra (a) részén látható) úgy definiálható, mint egy egy minta hosszúságú „be” jel, és kifejezhető:  $\delta(m) = \begin{cases} 1, & m = 0 \\ 0, & m \neq 0 \end{cases}$ , ahol az

$m$  változó jelöli a diszkrét idő mérőszámot. A Fourier transzformációt használva

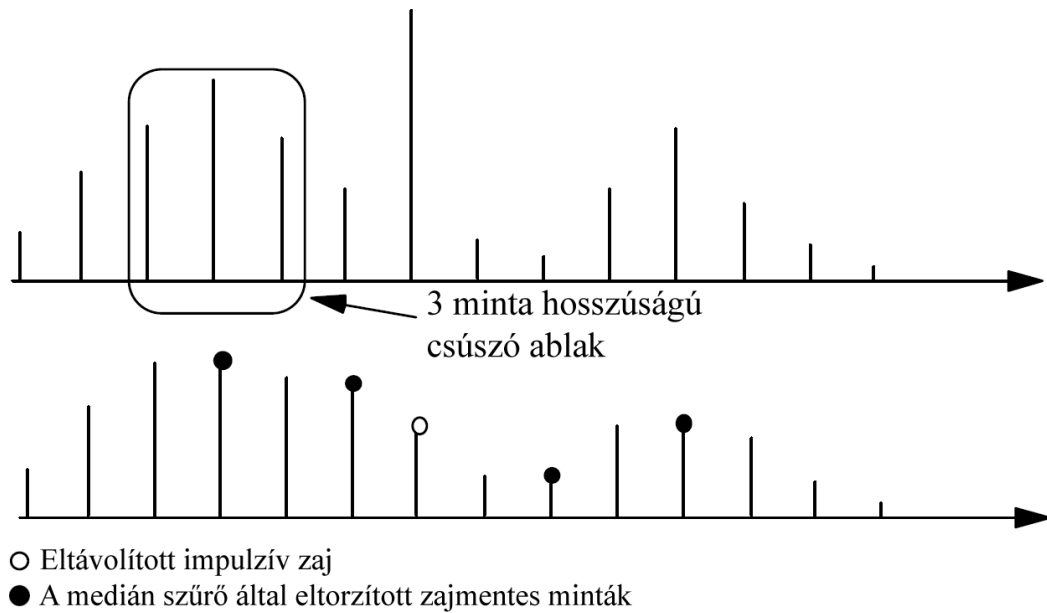
egy digitális impulzus frekvenciaspektruma megadható:  $\Delta(f) = \sum_{m=-\infty}^{\infty} \delta(m) e^{-j2\pi fm} = 1, 0,$

$-\infty < f < \infty$ .

Kommunikációs rendszerekbe a valódi impulzív típusú zajok hossza normálisan nagyobb, mint egy minta.

#### 3.3.4.2. A medián szűrő

Az impulzív zaj eltávolításának klasszikus megközelítése a medián szűrő. Medián alatt értjük az  $\{x(m)\}$  mintahalmaz elemeiből azt az  $x_{\text{med}}(m)$  értéket, melyre igaz, hogy a halmaz elemeinek fele ezen  $x_{\text{med}}(m)$  értéknél nagyobb, a másik fele pedig az  $x_{\text{med}}(m)$  értéknél kisebb. Tekintsük a mediánt minták egy halmazának, melyet úgy kapunk meg, hogy a mintákat rendezzük növekvő vagy csökkenő sorrendbe és kiválasztjuk a középső értéket. A medián szűrés esetén egy előre meghatározott hosszúságú ablakot csúsztatunk folyamatosan a jelen, és az ablakban található középső mintát kicseréljük az ablak mintáinak mediánjára, ahogy azt a következő ábra mutatja.



28. ábra Medián szűrő be- és kimenete. A működés közben torzítás is keletkezhet

A medián szűrő  $\hat{x}(m)$  kimenete az  $y(m)$  bemenet és egy  $2K+1$  mintát tartalmazó hosszúságú medián ablak alapján:

$$\hat{x}(m) = y_{\text{med}}(m) = \text{medián}[y(m-K), \dots, y(m), \dots, y(m+K)]$$

Egy számhalmaz mediánja a halmaz nem lineáris statisztikája egy nagyon jól hasznosítható tulajdonsággal, hogy nem érzékeny egy szokatlanul nagy mintaérték jelenlétére a halmazban, melyet kívülállónak hívunk. Ezzel szemben a középérték és részben a variancia érzékeny egy számhalmazban jelen levő kiugró értékre, vagyis az impulzív típusú zajra. Fontos tulajdonsága a medián szűrőnek – mely részben használható a képfeldolgozásban –, hogy megőrzi az éleket vagy a lépésenkénti folytonossághiányokat a jelben. A medián szűrők hatékonyan használhatók a képfeldolgozásban az impulzív zaj eltüntetésére az élek elmosása nélkül – ez nagyon jelentős a képfeldolgozásban. A medián filterrel történő kísérletezés során kiderült, hogy nem alkalmas magas minőségű eredmények produkálására audiojelek visszaállítása esetében. A medián szűrő nem tud mit kezdeni az olyan impulzív jellegű zajjal, mely hosszabb, mint egy vagy két minta. Továbbá a medián szűrő a hibásan impulzív zajnak vélt jelek módosításával jelentős mennyiségű feldolgozási torzítást okoz. A medián szűrő teljesítménye javítható egy alkalmazkodó határérték bevezetésével, tehát a mintát csak akkor cseréljük ki, ha annak eltérése a középső érték-

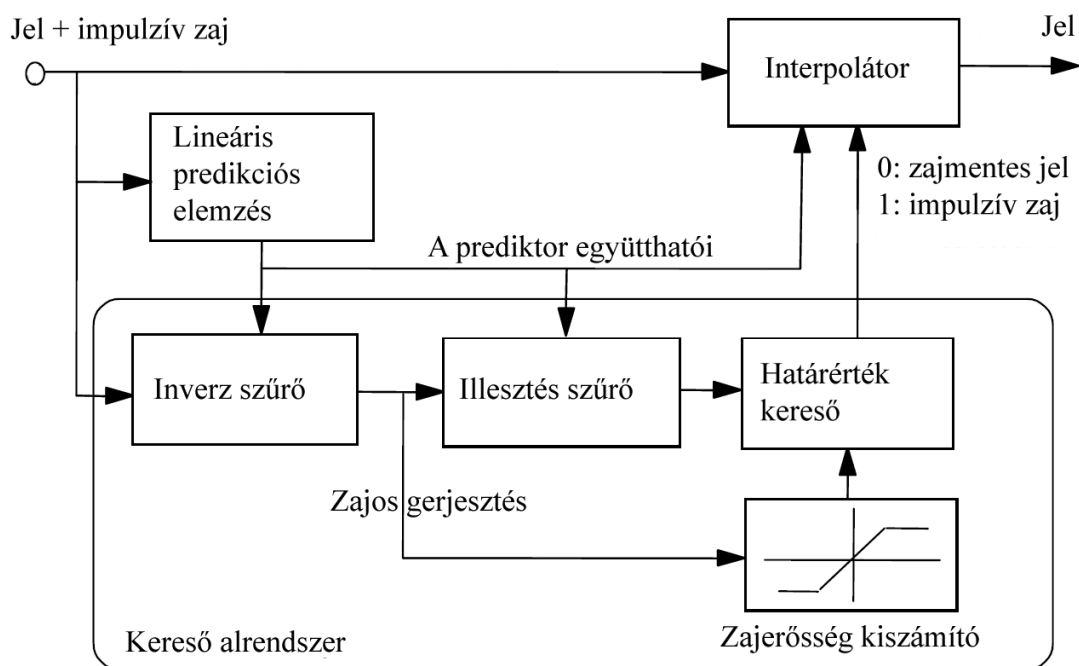
től meghaladja a határértéket:  $\hat{x}(m) = \begin{cases} y(m) & \text{ha } |y(m) - y_{\text{med}}(m)| < k\theta(m) \\ y_{\text{med}}(m) & \text{egyébként} \end{cases}$ , ahol

$\theta(m)$  egy alkalmazkodó határérték, mely erőteljes becsléssel határozza meg

$|y(m) - y_{\text{med}}(m)|$  átlagértékét,  $k$  pedig egy hangoló paraméter. A medián szűrők nem optimálisak, mert nem használhatók hatékonyan a jelgenerálás fiziológiájának előismerete vagy a jel és a zaj statisztikai eloszlásának modellezése nélkül.

### 3.3.4.3. Impulzív zaj eltávolítása lineáris prediktív modell segítségével

Az impulzív zavarok általában a jel egészét vizsgálva csupán csekély  $\alpha$  hányadát szennyezik. Mivel a jel jelentős  $1-\alpha$  része impulzív zajjal nem szennyezett, ezért előnyösebb megkeresni a zajos részeket (egyedülálló impulzusokat), és csak a torzult mintákat javítani. Ez a stratégia kiküszöböli a felesleges feldolgozásokat, és a minták egy viszonylag nagy része – mely nem szennyezett impulzív zajjal – érintetlenül eredeti állapotában marad. A következő ábra egy ilyen rendszer blokkdiagramját mutatja, mely két alrendszerből áll, egy keresőből és egy interpolátorból. A



29. ábra Egy impulzív zajeltávolító blokkdiagramja.

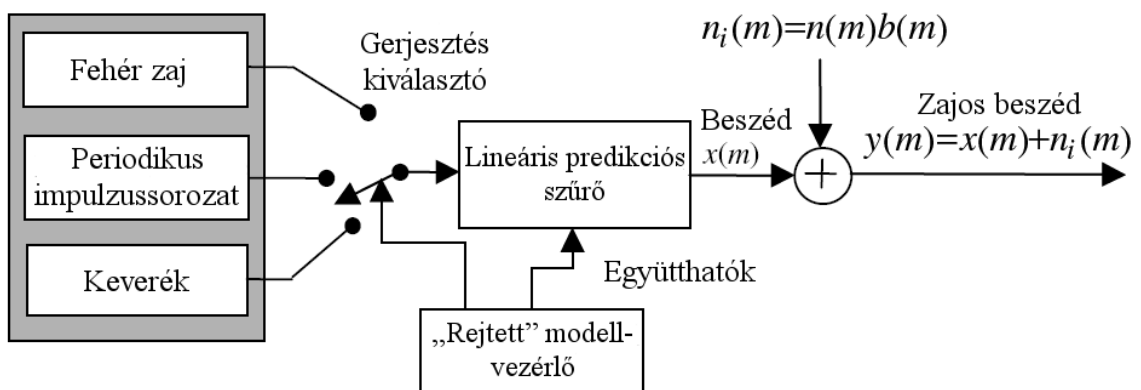
detektor megkeresi az impulzusok helyét, az interpolátor pedig kicseréli a torzult mintákat, figyelembe véve a torzult minta mindkét oldalán elhelyezkedő további mintákat. A detektor rész egy lineáris predikciós elemzőből, egy illeszkedésvizsgáló szűrőből és egy határérték keresőből áll. A kereső kimenete egy kétállású kapcsoló, mely az interpolátort vezérli. Impulzív zaj hiányában az interpolátor nincs bekapcsolva. Impulzív zaj jelenléte esetében az interpolátor bekapcsol és kicseréli a zaj által törölt mintákat.

## 3.3.4.3.1. Az impulzív zaj keresése

A legegyszerűbb módszer az ilyen jellegű zajok megtalálására, hogy meghatározzunk egy határértéket, és a határérték fölé eső mintákat zajnak osztályozzuk. Ez a módszer tökéletesen működik nagyon nagy impulzusok esetében, de ha a zaj amplitúdója a megadott határérték alá esik, akkor nem találja meg. A keresés eredményesebbé tehető, ha kihasználjuk a jel és az impulzív zaj jellemzőinek különbségét. Az impulzív zaj, vagy rövid időtartamú hullámvész bevisz egy a jelre nem jellemző megszakítást a jelbe. A megszakítás könnyebben megtalálható, ha a jelet szétválasztjuk. A szétválasztás (vagy digitális jelnél megkülönböztetés) művelete egyenértékű a kölcsönösségi viszony megszüntetésével, vagy a spektrális fehérítéssel. A lineáris prediktor modellezi a jel kölcsönösségi struktúráját, és ezen összefüggések megszüntetését végzi az inverz szűrő. A zajmentes  $x(m)$  jel lineáris predikciós modell segítségével a következő összefüggéssel írható fel:

$$x(m) = \sum_{k=1}^P a_k x(m-k) + e(m), \text{ ahol } a = [a_1, a_2, \dots, a_P]^T \text{ a lineáris prediktor } P \text{ rendű}$$

együttható vektora, és az  $e(m)$  gerjesztés egy zajhoz hasonló jel, vagy vegyes véletlenszerű zaj keveréke, és hullámvészok kvázi periodikus sorozata, melyet a következő ábra mutat. Az impulzív zajkereső nem a korrelálatlan kétállapotú impulzív zaj



30. ábra Zajos beszédmodell. A jelet egy lineáris prediktor modellezi, az impulzív zaj modellezésére pedig egy kétállapotú amplitúdómodulált modellt használunk.

alapján dolgozik, hanem azon a megfigyelésen alapszik, hogy a lineáris prediktorok jó modellt biztosítanak korrelált jelekre. A zajos  $y(m)$  jel a prediktor gerjesztett jelévé történő alakítása a következő hatásokat váltja ki:

- a.) Az jel amplitúdójának léptéke csökken az eredetileg gerjesztett jeléhez képest, mivel a zaj amplitúdójának léptéke változatlan marad, vagy növekszik.

- b.) A jel nem lesz kölcsönösen összefüggő, mert az impulzív zaj elmaszatólódott és átalakult léptékes alakba az inverz szűrő impulzus válaszaként.

Mindkét hatás javítja az impulzív zaj megtalálhatóságát. A beszéd és a zene hang időtartamának rezonanciája vagy a hangszer által erősített és spektrálisan alakított gerjesztésekből áll össze. Ez a gerjesztés véletlenszerűbb a beszédnél, és gyakran sokkal kisebb az amplitúdó-tartománya. Az impulzív zaj megtalálhatósága javítható az inverz szűréssel, mely lényeges lehet és függ a jel időben változó korrelációs struktúrájától. Ez a módszer egy optimális határértékképző használatával képes hatékonyan csökkenteni az impulzív zaj keresésének problémáját, a sorból kilógó minták megtalálását.

#### 3.3.4.3.2. Az impulzív zaj megtalálhatóság-javításának elemzése

A következőkben az impulzív zaj megtalálhatóságának javítását elemezzük ez inverz szűrő kimenetének vizsgálatával. Az előző egyenlet használatával a zajos jel modelljét a következő egyenlet segítségével (az előző átírt változata) írhatjuk fel:

$$y(m) = x(m) + n_i(m) = \sum_{k=1}^P a_k x(m-k) + e(m) + n_i(m),$$

ahol  $x(m)$ ,  $y(m)$  és  $n_i(m)$  rendre a

zajos jel, a jel és a zaj. Egy számítást használva  $\hat{\mathbf{a}}$  lesz  $\mathbf{a}$  prediktor együttható vektora, és a zajos  $y(m)$  jel inverz szűrhető és átalakítható  $v(m)$  zajosan gerjesztett jellé így:  $v(m) = y(m) - \sum_{k=1}^P \hat{a}_k y(m-k) = x(m) - n_i(m) - \sum_{k=1}^P (a_k - \tilde{a}_k)[x(m-k) + n_i(m-k)]$ ,

ahol  $\tilde{a}_k$  a prediktor együttható számítási hibája. A zajos jel modelljének egyenlete ez alapján az egyenlet alapján következő egyenleté írható át egy másik alakba:

$$v(m) = e(m) + n_i(m) + \sum_{k=1}^P \tilde{a}_k x(m-k) - \sum_{k=1}^P \hat{a}_k n_i(m-k).$$

Ebből az egyenletből három tétel határozható meg, mely elősegíti a zaj megtalálását a gerjesztési szekvenciában.

- a.) az  $n_i(m)$  impulzív zavar, mely általában domináns;
- b.) a régebbi  $P$  darab impulzív zajminta hatása az inverz szűrő működése végett, a jelenbe maszátolva:  $\sum_{k=1}^P \hat{a}_k n_i(m-k)$ ;
- c.) a gerjesztett jel varianciájának növekedése a paramétervektor számítási hibái miatt, mely a következő kifejezéssel írható le:  $\sum_{k=1}^P \tilde{a}_k x(m-k)$

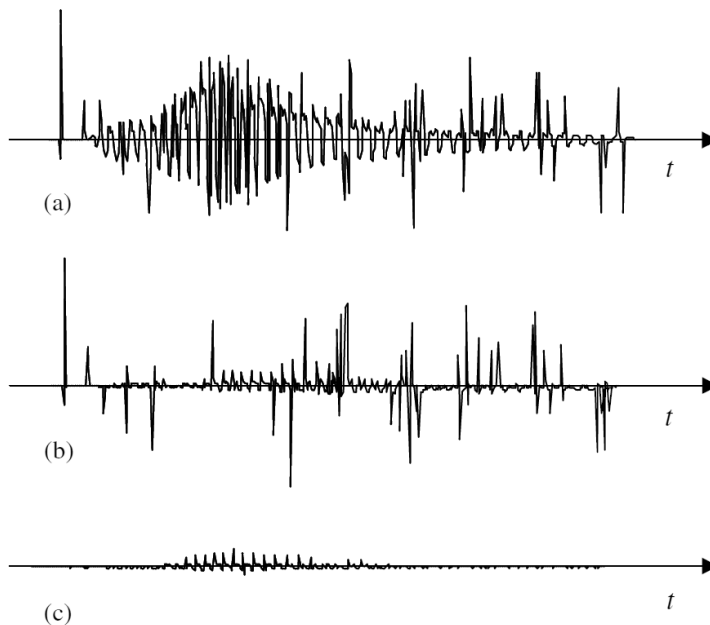
Az inverz szűrő eredményeinek javítása a következő formulákkal írható le. Az impulzív zaj a zajos jelhez viszonyított aránya az alábbiakban ismertetett módon

adható meg:  $\frac{\text{impulzív zaj energiája}}{\text{jel energiája}} = \frac{E[n_i^2(m)]}{E[x^2(m)]}$ , ahol  $E[\cdot]$  a várható érték operátora.

Az impulzív zaj detektálásában az érdeklődés tárgyát képező jel az impulzív zaj lesz, melyet a kísérő jelből kell megtalálni. Feltesszük, hogy a domináns zaj a  $v(m)$  zajos gerjesztési jelben az  $n_i(m)$  impulzív zaj, a két jel aránya a következőképpen írható fel:  $\frac{\text{impulzív zaj energiája}}{\text{gerjesztési energia}} = \frac{E[n_i^2(m)]}{E[e^2(m)]}$ . A teljes impulzív zaj jellel szembeni

aránya az előző két egyenlet hányadosából kapható meg, melyet a következő egyenlettel tudunk leírni:  $\frac{E[x^2(m)]}{E[e^2(m)]} = \text{nagyobbodás}$ .

Ez az egyszerű elemzés megmutatta, hogy az impulzív zaj megtalálhatósága függ a teljesítményerősítés karakterisztikájától, ennek következtében a lineáris predikciós modell rezonanciáitól. Beszéd esetében a jel amplitúdója zajmentes gerjesztett jelről lévén szó  $10^{-1}$  és  $10^{-4}$  között változik, ami a beszéd maga, ezért jelentős javulás érhető el az impulzív zaj megtalálhatóságában az inverz szűrő alkalmazásával a zajos beszédre. A következő ábra az impulzív zaj megtalálhatóságának



31. ábra Az inverz szűrő hatása a impulzív zaj megtalálhatóságára. (a) impulzív zajjal szennyezett beszéd, (b) az impulzív zajjal szennyezett beszéd gerjesztése, (c) beszéd gerjesztés impulzusmentesítve.

javítására alkalmazott inverz szűrő hatását mutatja. Az inverz szűrés hatása, hogy az  $x(m)$  jelet korrelálatlan  $e(m)$  jelgerjesztéssé alakítja, míg az impulzív zaj az inverz szűrő impulzív  $[1, -a_1, \dots, -a_p]$  zajátvitelének megfelelően torzul és megnövekszik, ahogy azt az iménti tétel mutatja a következő összefüggés alapján:

$\sum \hat{a}_k n_i(m-k)$ . Feltételezzük, hogy a gerjesztés egy

normális eloszlást követő fehér zaj, egy szűrő, melynek együtthatói az inverz szűrő együtthatóival egyeznek képes növelni a torzult impulzív zaj megtalálhatóságát a gerjesztési jelben.

## 3.3.4.3.3. Kétoldali prediktor az impulzív zaj megtalálására

Az impulzív zaj megtalálásának egyik hatékony módja, hogy a jelet korrelálatlanná tesszük. Ez a módszer felhasználható a kétoldali lineáris predikciós modellben is. Egy  $x(m)$  minta kétoldali lineáris predikciója  $P$  darab régebbi (múltbeli) és  $P$  darab jövőbeli minta alapján dolgozik, és a következő egyenlettel definiálható:

$$x(m) = \sum_{k=1}^P a_k x(m-k) + \sum_{k=1}^P a_{k+P} x(m+k) + e(m),$$

ahol  $a_k$  értékek a kétoldali prediktor együtthatói, és  $e(m)$  a gerjesztési jel. Minden elemzés arra az esetre készül, mikor csak egy oldalról vizsgáljuk a jelet, de ez kiterjeszthető a kétoldali modellre is. Azonban a gerjesztés mértéke a kétoldali modell esetében kisebb varianciát mutat, mint az egyoldali modellben, mivel az imént leírt egyenletben mind a múltbeli és mind a jövőbeli összes mintára modellezett a minták korrelációja. Habár az iménti egyenletben egy nem véletlenszerű szűrő van, az inverze szükséges a kereső alrendszerbe, mely már véletlenszerű. A kétoldali prediktor használata további javulást eredményezhet az impulzív zaj megtalálhatóságában.

## 3.3.4.3.4. Az eldobott minták interpolációja

Az impulzív zaj hatására a minták visszavonhatatlanul torzulnak, és eltávolításukkal „űr” keletkezik, melyet interpolációval próbálhatunk kitölteni. A hiányosságok interpolálására – ahhoz, hogy ne maradjon hallható nyomuk –, magas hanghűségű interpolátor szükséges. A legkisebb négyzetes autoregresszív (LSAR) módszer magas minőséget eredményez viszonylag kis számú az impulzív zaj miatt elveszett minta esetében. Az LSAR interpolációs módszer egy kétszintű folyamat. Az első szinten az impulzív zaj mindkét oldalán megmaradt minták alapján becsléssel meghatározzuk a jel predikciós modelljének paramétereit. A második szinten a meghatározott modellparaméterek és a hézag két oldalán található minták alapján interpoláljuk az elveszett mintákat. Ezzel az interpolációs módszerrel az impulzív zaj által eltorzított audiojelek magas minőségben visszaállíthatók.

## 4. Zajsűrés a gyakorlatban

### 4.1. Zenekari stúdiómunka

#### 4.1.1. A felvétel

Egy hangstúdióban legyen szó akár könnyűzenei vagy komolyzenei műről a hangrögzítés során számos problémával kell szembenéznünk. Amennyiben megoldható, az akusztikus (nem elektromos) hangszerek felvételét úgynevezett süket-szobában célszerű kivitelezni, a visszhanghatások elkerülése végett, melyek a hanganyag torzulását okozhatják. A felvételt adott esetben csatornákra bontva külön-külön rögzítjük. Ez azt jelenti, hogy az egyes hangszerek külön játsszák a darabot, majd a felvételt követően egy keverőpult segítségével születik meg a végleges hanganyag.

A felvétel során is felütik a fejüket a zajok. Amivel mindenképpen számolni kell, hogy a bemikrofonozott jelforrás valamilyen távolságra helyezkedik el a mikrofontól, néha a zenész átéléssel játszik, és a hangszerrel együtt mozog. A távolság és az alkalmazott mikrofon minőségének függvényében a rögzítő technika mindenképpen rögzít egy kevés környezeti zajt is. Ezen zajok ellen a környezet megfelelő megválasztásával lehet a legkönnyebben védekezni.

További problémákat okoznak a mikrofonok viszonylag alacsony jelszintjei, és a jeleket továbbító kábelek. A kábeleket célszerű lehetőség szerint minél vastagabbra, tiszta anyagból készültre, és minél rövidebbre választani, a lehető legkevesebb csatlakozóponttal, ugyanis ez tudja a jel legjobb minőségű szállítását elvégezni. A nagy keresztmetszet és a viszonylag kis hossz biztosítja a kis ellenállást, mely a jelveszteséget csökkenti, és a tiszta anyag (például oxigénmentes réz, négy kilences tisztasággal – 99,99 % réz) pedig a szennyezések miatt kialakuló örvényáramok okozta zajtól óvnak bennünket. A csatlakozópontok további torzítást és zajt eredményezhetnek, az anyagok különbözősége, és a csatlakozó felületek mérete miatt. További eredményjavulás várható az arany, vagy felületén aranyozott kábelek alkalmazásától a magas frekvenciájú jelek átvitele szempontjából – a skin effektus miatt. A skin effektus azt jelenti, hogy a váltakozó áramot szállító vezetékben az elemi töltéshordozók a vezetőben annál távolabb haladnak a vezető középpontjától, minél nagyobb az átvitt jel frekvenciája.



Az alacsony jelszint miatt a digitális feldolgozhatóság elérése érdekében előerősítő fokozatokat kell a rögzítő készülék és a jelforrás(ok) közé építeni, melyek esetében az alkatrészek kialakítástól és minőségétől függően termikus zajgenerátorként működnek. A sávátviteli problémákat többnyire nem az erősítőelemek okozzák, hanem a mikrofonok.

További érdekesség a jelforrás szempontjából, hogy az elektromos hangszereket, mint például egy elektromos gitár, vagy basszusgitár általában egy úgynevezett alap erősítőre csatlakoztatjuk, majd ezen erősítő hangszórója elé helyezünk el mikrofont. Általában e hangszerek esetében az effekteket már az erősítő előtt, vagy annak effektútjába csatlakoztatjuk. Ilyenkor egy torzító effekt képes a hangszer hangszedője (pickupja) által felfogott zajokat nagymértékben felerősíteni, amit még – egy a torzító előtt elhelyezett (a játéktechnikát könnyítő) – dinamika kompresszorral lehet „növelni”. Ezek a zajok elektromágneses zajok, és sajnos a hangszer működéséből adódóan nehéz ellenük védekezni, mivel a húrok rezgését is ilyen módon érzékeli a hangszedő. Ilyen hangszerek esetén már két termikus zajgenerátor rendszer kapcsolódik a rögzítő elé, ezzel fokozva a zajszennyezést.

A rögzítő technika manapság már többnyire digitális rendszer, melynek az érkező jeleket egy ADC segítségével digitálissá kell alakítania. A jel ekkorra már zajjal erősen szennyezett lesz, a korábbi berendezések felépítésének és minőségének függvényében. A digitalizálás a mai korszerű hangkártyáknak nem jelentenek problémát, mégis stúdió célokra külön gyártanak – magasabb árfekvésű, jobb minőségű – kártyákat. A mintavételezési frekvencia megválasztása ekkor az egyik fő probléma. Általában a végeredmény CD lemezre kerül, ahol szabvány írja elő a 44100 Hz-es mintavételezési frekvenciát. Vannak hangkártyák, melyek ilyen mintavételezési frekvencia mellett nem produkálnak lineáris sávátvitelt, ellenben 48000 Hz-en igen. A felvett jelet visszahallgatva, van különbség 48000 Hz-en és a 44100 Hz-en mintavételezett jelek között az első javára, bár azt gondolhatnánk, hogy a hallható tartományon már ezek a frekvenciák kívül esnek. Másik érdekes kérdés a kvantálási hossz. A mai világban már a megfelelő dinamika elérése érdekében 16 bitnél nagyobb értéket választunk.

#### 4.1.2. Az utómunka

Az utómunkálatok első lépése a felvétel során a jelre került zajok eltávolítása. Ez esetben, ha sávonként rögzítettünk, akkor minden sávra külön célszerű a zajt eltávolítani, egészen pontosan annak hatását csökkenteni, ugyanis a zajok vélet-

lenszerű mivolta nem teszi lehetővé teljes eltávolításukat. Az ilyen felvételek esetén azért kell külön kezelni a zajokat, mert egy mikrofon és egy elektromágneses hangszedő több zajgenerátorral együttműködve más és más jellegű zajokat eredményeznek. Ha például egy dobfelszerelésre gondolunk, akkor észrevehető, hogy a megfelelő minőségű jel érdekében célszerű a hangszer egyes darabjait külön be-mikrofonozni, adott esetben más típusú mikrofonokkal, tehát ekkor már egy mikrofonrendszerrel beszélünk. Az egyes mikrofonok hangját külön kell bontani hangszín szempontjából, esetleg sztereó felvétel esetében még a panoráma sem fog egyezni minden esetben (például: lábcintányér jobboldal, pergődob közép, kísérő cintányér baloldal, természetesen a zenész kezességét figyelembe véve lehet fordítva is). A hangszínt minden esetben az egyes részeken külön kell beállítani, ami függ a mikrofon minőségétől, távolságától, helyzetétől és állásától. Ez a mikrofonrendszer mikrofononkénti hangszín beállításokkal másfajta zajt állít elő, mint például az imént említett elektromos gitár hangszedője egy torzítóval. Ezért is fontos a sávonkénti felvétel. Ezek a zajok csak ritka esetben tartalmaznak impulzív jellegű zajokat (vagy ha az ilyen zajok megjelennek, azok megszüntetéséről még a felvétel előtt célszerű gondoskodni), ezért a zaj csökkentésére a legalkalmasabb módszer a spektrális kivonás. Fontos, hogy a zajszűrő program (algoritmus) olyan legyen, hogy a hangképet ne, vagy csak nagyon kis mértékben módosítsa, a legjobb eredmény elérése érdekében.

Ezt követően az egyes hangszerek még további effektek hozzáadásával színesíthetők, valamint itt nyílik lehetőség az utolsó hangszínbeállításra, ugyanis a felvételkor még csak nyersanyag keletkezik. Itt adott esetben az egyes sávokat további analóg keverőpultba küldhetjük, de léteznek olyan stúdióban alkalmazható programok, melyek képesek a többsávós felvétel kezelésére. Ilyen program például a korábbi (a 2003-ban az Adobe cég által megvásárolt Syntrillium cég fejlesztésében megjelent) Cool Edit, és a manapság helyette alkalmazott Adobe Audition program, melyek kereskedelmi forgalomban kaphatók, de van ingyenes időkorlátos próbaváltozatuk is. Ezek a programok nem csak többsávós hanganyag effektezésére, hangszínbeállítására alkalmasak, hanem az azt megelőző zajcsökkentésekre is.

Természetesen a megfelelő hangzás elérése érdekében a hanganyagot a hangfelvétel elkészülte után hetekkel célszerű keverni, majd a végeredményt többféle berendezésen visszahallgatni, a maximális minőség elérése érdekében. A megfelelő hangzás elérése után még egy dinamika kompresszor segítségével lehet az anyagot

„összetömöríteni”. A mintavételezési frekvencia és a kvantálási hossz a későbbi hanghordozó függvényében változhat a rögzítési értékekhez képest, de ez sajnos a jel torzulását és néha még további zajok megjelenését is okozhatja.

## **4.2. Hagyományos hanglemez, hangkazetta archiválása CD-re**

Minden archiválási folyamat az adott hanghordozón tártolt hanganyag digitalizálásával kezdődik. Ezt akár otthonunkban is megtehetjük, feltéve, hogy rendelkezünk a forrás hanghordozó megfelelő minőségű lejátszására alkalmas berendezéssel. A hangminőséget jelentősen befolyásolhatja a jelforrás, továbbá az is, hogy milyen eszközzel végezzük a digitalizálást, és hogy milyen kábelek kötik össze az adott berendezéseket. Itt is lehet környezeti elektromágneses zavarokra számítani.

### *4.2.1. Hanglemez archiválása*

#### 4.2.1.1. A lejátszó rendszer

A gyártók általában többféle megoldást alkalmaznak a lemezejátszó mechanikájának és a mechanikai mozgásokat elektromos jellé alakító részegységek kivitelezésére.

A mechanika leglényegesebb elemei:

- Lemeztányér, amely a lemez forgatásáért, megtartásáért felelős, továbbá sok esetben a hangkar mozgatásával is foglalkozik, az alsó mechanikus szerkezeten keresztül. Többnyire nagy tömegű, az egyenletes forgás megvalósítása érdekében (nagyobb tehetetlenség, így nem képes a motor gyors fordulatszám-változásainak követésére). Fontos a jó csapágyazás és kenés a könnyű futás érdekében.
- Motor, amely a lemeztányért hajtja, nagyon fontos, hogy állandó fordulatszámú legyen nagyon kicsi ingadozással, mely valamilyen általában erőzárás közlőművel kapcsolódik a lemeztányérhoz (ez esetben fordulatszáma sokkal nagyobb, mint a tányéré).
- Hangkar (esetleg a mozgató mechanikával együtt), mely hangszedő fejjel együtt a lemeztűt egy shellnek nevezett tartóban rögzíti, és gondoskodik a megfelelő tűnyomásról és a barázdakövetésről.

A lemeztányér meghajtására háromféle megoldást szoktak alkalmazni, az egyik a dörzshajtás, a másik a szíjhajtás – gyakorlatilag ez az egyik legelterjedtebb –, és végül a direkt hajtás – ez közvetlen a motor forgórészére rögzített viszonylag

könnyű lemeztányért jelent. A motor lehet váltakozó áramú, mely többnyire a szabványos hálózatban alkalmazott frekvenciára épít a forgás tekintetében, vagyis szinkron motor, és lehet úgynevezett DC (egyenáramú) motor is. A közvetlen hajtású lemezjátszók esetében a legpontosabb fordulatszám elérése végett kvarcvezérlést alkalmaznak. Egyes típusoknál egy stroboszkóp lámpa segíti a fordulatszám beállítását, egy potenciométer beiktatásával, melyet általában az egyenáramú motorral szerelt típusok esetében alkalmaznak. A három megoldás közül a kvarcvezérlésű direkt hajtás a legpontosabb megoldás.

A hangkart mechanikusan áthelyezni képes lemezjátszók a lemeztányér alatt elhelyezett mechanikus szerkezet segítségével teszik mindezt. Létezik belőlük félautomata és automata kivitel is. A hangkarban helyet foglaló (sok esetben leszerelhető, cserélhető) shell biztosítja a lemeztű számára kialakított hangszedő fej rögzítését. A szögkiegyenlítés végett – nem gyakran ugyan – alkalmaznak úgynevezett tangenciális hangkart, melynél nincs ilyen jellegű probléma. A nem tangenciális esetben a fej valamekkora szöget zár be a hangkarral, melyet vagy a hangkar meghajlításával, vagy a shell kialakításával érnek el.

Hangszedőből háromféle alaptípus létezik, melyek működésükben különböznek egymástól.

Az egyik – régebben viszonylag olcsó és hatékony megoldást biztosító – típus a kristály hangszedő. Ez működését tekintve a gyémánt vagy zafír tű által a lemeztől letapogatott rezgéseket egy vékony cső segítségével egy piezokristályra vezeti, mely a rezgést elektromos jellé alakítja. Ez a letapogató rendszer viszonylag nagy túnyomással működik jól, és ez sajnos nem kíméli a mikrobarázdás lemezünket. Hangzás tekintetében nem túl jó, de kielégítő eredmény elérhető vele. Nagy előnye, hogy nem igényel külön korrekciós előerősítő fokozatot, viszont impedanciaillesztés szempontjából nagyimpedanciás bemenetet igen.

A másik szinte a legelterjedtebb, már jó hangzást biztosító megoldás a mozgó mágneses hangszedő. Ez esetben (sztereó kialakítást feltételezve) két egymással 90°-os szöget bezáró tekerccs között mozgatja a vékony cső végére rögzített mágnest a lemeztől letapogatott mechanikai rezgés. Túnyomás tekintetében változatos megoldásokkal találkozhatunk, mert van itt  $\frac{3}{4}$  grammtól kezdődően kristály hangszedőt megszügyenítő 5 grammos nyomást igénylő kialakítás is. Itt már kell egy – a köznapi életben alkalmazott erősítők nagytöbbségében az alkalmazás idő intervaluma alatt fellelhető – korrekciós előerősítő fokozat, mely a mély hangokat kiemeli, a közép hangokkal nem tesz semmit, és a magas hangokat pedig csillapítja. Erre

azért van szükség, mert a lemez előállításakor így rögzítik a jelet – a barázda szélességének csökkentése miatt (több hanganyag fér el egy lemezen). Viszonylag egyszerűen cserélhető az elkopott tű, melyet még manapság is forgalmaznak bizonyos helyeken.

A harmadik megoldás a mozgó tekercses megoldás. Ez csak fejjel együtt cserélhető, nagyon drága, és hangzásvilágát tekintve ez a legjobb minőségű. Az előerősítő fokozat ide is kell, de az erősítésének nagyobbak kell lennie az előzónél, mert a jel feszültsége csupán kb. az egytizede a mozgó mágneses hangszedőének. Létezik olyan erősítő, amely mindkét jelet képes fogadni. A megoldások közül a legjobb hangminőség a mozgó tekercses hangszedővel érhető el, de ára, és a szükséges további eszközök miatt inkább a mozgó mágneses hangszedőt használjuk.

#### 4.2.1.2. Várható zajok

A hanglemezek mechanikus adathordozók, ezért nagyon sérülékenyek. A nem megfelelő tárolás következtében a por komoly károkat tud okozni a hangzásban. A nem megfelelő poreltávolítás apró karcokat okoz, melyek hallhatók lesznek lejátszáskor. A lemez felületére az ujjainkról került zsír is károsítja a hanghordozót, továbbá a felületen nagyobb karcok is kialakulhatnak (a nem megfelelően mozgatott hangkar következtében a tű is okozhat ilyen karcokat), melyek az impulzív zajokért felelősek.

A hanglemez kialakítása során a felületi egyenetlenségek keletkeznek, melyek szintén hallhatóvá válhatnak lejátszáskor. Ezért van különbség a keleti és a nyugati nyomású lemezek között, mivel más a gyártás alapanyaga és némileg a technológiája is.

A lemezejátszó mechanikája is képes zajokat előállítani. A lemeztányér a csapágyazás és kenés függvényében valamilyen mértékben rezeg. Ez egy precízen csapágyazott motor esetén egy direkt hajtású lemezejátszóban nem okoz problémát, mert általában a hallható tartományon kívül esik a rezgés frekvenciája.

Különbséget kell tenni a lemez száraz, illetve nedves lejátszása között is. A nedves lejátszás esetén a lemezt mosófolyadékkal lejátszás közben mossuk, így a szennyeződések eltávolítása is megtörténik, továbbá a folyadék vékony réteget képez a tű és a barázda információt tartalmazó részei között, ezért a hangzás is változhat nagyon kis mértékben.

A lemez kivitelezésénél fogva a korrekciós előerősítő fokozat is zajt termel, majd az azt követő illesztő fokozat viselkedik zajgenerátorként. Ezekben az erősí-

tőkben az aktív elemek melegszenek, termikus zajt előállítva. Továbbá a nagy erősítés és a mágneses működés miatt érzékenyek az elektromágneses zajokra is, mely megfelelő szigeteléssel orvosolható.

Tehát a lemezen megjelenik a helyreállítandó hézagokat maga után vonó impulzív zaj, és a spektrumát tekintve nemfehér zaj is, melyeket két különböző eljárással tudunk csökkenteni.

#### 4.2.1.3. Zajcsökkentési eljárás

A felvétel elkészítésekor figyelni kell az alkalmazott technikák jelszint és impedancia-illesztésére, valamint a mintavételezési frekvenciára és a kvantálási hossza is. Lehetőleg a cél adathordozóéval kompatibilis értékeket kell választani, a minőségromlás elkerülése érdekében.

A lemezjátszók készülhettek a keleti, illetve nyugati régióban is. A keleti régió általában az ötpólusú tuchel aljzatokat alkalmazta a jelek továbbítására, amely nem csak kialakításában, hanem jelszintjében és impedanciájában is eltér a nyugati régióban elterjedt (4 darab) RCA csatlakozókhöz képest. A tuchel csatlakozók kisebb jelszinten dolgoznak, és nagyobb bemeneti impedanciát igényelnek, továbbá kisebb felületen érintkeznek, tehát nem alkalmasak közvetlenül digitalizálásra. (Természetesen egy egyszerű erősítőfokozat közbeiktatásával alkalmassá tehetők a feladat elvégzésére az ilyen berendezések is.) A célszerű választás tehát a folyamat kivitelezésére egy *nyugati gyártású lemezjátszó, mozgó mágneses* hangszedővel, a megfelelő előerősítő fokozatot tartalmazó szintén *nyugati gyártású erősítővel*, mely illeszthető a szabványos hangkártyák vonalbemenetére.

A folyamat a digitalizálással kezdődik. A feladatra a számítógépbe célszerű egy eleve jó értékű jel–zaj aránnyal rendelkező, kis torzítású hangkártyát választani. A Creative cég régebben is készített ilyeneket, és ma is gyárt alkalmas eszközöket. (A Sound Blaster sorozat elemei.) A digitalizáláskor a jel erős zajszenyezéssel érkezik az eszközbe.

Az első lépés az impulzív zajok megkeresése és eltávolítása. Erre alkalmas az előző részben említett Adobe Audition program. Itt a határértékek feltérképezése után indulhat a folyamat, sajnos az eredmény csak a végrehajtás után lesz hallgatható, nincs „előzetes bemutatás” lehetőség. A folyamat az impulzív zaj mennyiségétől és a számítógép sebességétől függően akár több órát is igénybe vehet. Közben látható, hogy hány mintát vélt zajnak, és hány mintát sikerült kijavítania a programnak. Természetesen az javított/kihagyott arány a beállítás függvénye, na-

gyon karcos lemez esetén sok impulzív zaj van, melyek javítása eredményezheti a hangminőség romlását is, oly formában, hogy a magas frekvenciájú jeleket is impulzív zajnak tekintik.

A következő lépés a spektrális kivonás. Ekkor mintát kell venni egy csak zajos részből, mely lehet a lemez bevezető része. A mintában szereplő pillanatképek száma beállítható, de nem feltétlenül eredményez jobb minőséget a nagyobb szám. Fontos paraméter a zaj csökkentésének mértéke, a precizitási szorzó és a simítás is. Ezek együttesen határozzák meg az eredmény minőségét. Van „előzetes bemutatás”, tehát meghallgathatjuk, hogy milyen eredményre számíthatunk, sőt megfigyelhető csak a zaj is.

A végén célszerű a jelet normalizálni, a hangerő (vagy jel–zaj arány) javítása érdekében. A mentés történhet szakaszonként is, ezzel a hanganyag darabolása is megtörténhet. Az eljárással akár 60–80 dB jel–zaj arány is elérhető, ami a kiindulási 30–40 dB-hez képest nem rossz eredmény.

#### 4.2.2. Hangkazetta archiválása

##### 4.2.2.1. A magnetofon és a hangszalag

A magnetofon vékony mágnesezhető réteggel bevont műanyag szalagon rögzíti a jeleket. A jelek, mivel az átmagnesezhető réteg vékony, nagyon kis feszültséget indukálnak a letapogató rendszerben, vagyis a magnetofon fejben, mely egy speciális kemény, kopásálló bevonattal rendelkező tekercs(pár). A kicsiny indukált feszültség erősítése egy korrekciós előerősítő fokozattal kezdődik, melynek a magas hangokat kell kiemelnie. A lejátszó rendszer ezt a szalagot húzza egyenletes sebességgel a fej előtt. A szalagsebességről egy kisebb átmérőjű fém és egy rugóval rászorított nagyobb átmérőjű gumigörgő gondoskodik, egyes esetekben ezt a rendszert megduplázzák, a stabilabb szalagsebesség érdekében. A görgők viszonylag könnyen szennyeződnek, ami a szalag nem kívánt oldalirányú mozgásához vezet; ez károsan befolyásolja a lejátszást. A szalagsebesség egy fontos jellemzője a berendezésnek, ami 4,76 cm/s egy átlagos magnetofon esetében. Ennek beállítása referenciaszalaggal és oszcilloszkóppal történik. A gyártók egy túrésen belülre állítják ezt az értéket, tehát két azonos szalagsebességű berendezést nehéz találni. A lejátszó és felvevő kombinált fej szennyeződhet, átmagneseződhet, ami szintén minőségromlást eredményezhet. Fontos paraméter még a szalagon található sáv és a fejben található érzékelő egymáshoz viszonyított elhelyezése is. Nem megfelelő beállítással a magasabb frekvenciájú hangok mennyiségének nagymértékű csökkenése

érhető el. A gyártók ezt is határértékek közé állítják, de létezik gyárilag rosszul beállított magnetofon is.

A szalag esetén problémák a szalagra felvitt mágnesezhető réteg minőségével kezdődnek, ebből alapjában véve négyféle típust alkalmaztak. Az első a normál vasoxid, a második a króm dioxid, a harmadik a ferrokróm (ezt már nem alkalmazzák) a negyedik pedig a magnetit. A következő probléma a mágnesezhető réteg szemcsézete, egészen pontosan a mágneses információhordozók elrendezése. Ez a szalag gyártójától függően – azonos típusú mágnesezhető anyag esetén – anyagminőségben is eltérhet egymástól, tehát vannak jobb minőségű és rosszabb minőségű hangkazetták, annak függvényében, hogy a mágnesességet mennyi ideig képesek megőrizni az egyes szalagok. A harmadik probléma a műanyag réteg vastagsága, mely lejátszáskor is okozhat problémát a szalagsebesség és az oldalirányú mozgás miatt, de tárolás esetén kérdéses, hogy az egyes átmágnesezett rétegek hogyan hatnak egymásra az idő múlásával. A műanyag hosszirányban képes a megnyúlásra, ami tovább fokozza a problémákat. Továbbá a vékony szalag nagyon érzékeny a mechanikai behatásokra, mint például a gyűrődés, amely durva minőségromláshoz vezethet. Az ilyen jellegű hiányok sem pótolhatók a jel folytonosságában, még digitális módszerrel sem. Nagyon fontos szerepe van még szalagok esetén az előmágnesezésnek, ugyanis ez is potenciális zajforrás.

#### 4.2.2.2. Várható zajok

A szalagos egységek esetében az előmágnesezési zaj után a szalagnak is van egy meghatározott értékű zaja, melyet egy olyan elektronika követ, ahol több termikus zajgenerátor is működik egy időben. A zaj ezért összetett, és általában egy elég jól hallható sistergés formájában jelentkezik, tehát főleg a magasabb frekvenciákon ölt jelentős mértéket.

A magnetofon esetén alkalmazható a Dolby zajcsökkentő rendszer, de lejátszáskor csak akkor, ha a felvétel is azzal készült, ellenkező esetben nem csak a zaj, hanem a magas hangok is eltűnnek. Ha nem tudjuk egy szalag esetén, hogy használtak-e valamilyen analóg zajcsökkentőt a felvételkor, akkor a berendezés zajcsökkentőjét kapcsoljuk ki, mert így a digitális zajszűrő hatékonyabban távolíthatja el a zajt, és több hasznos jel maradhat épen.

A magnetofonnal – a gyártók ígéretei szerint – akár 60 dB jel–zaj arány is elérhető a Dolby zajszűrővel, de sajnos a szalagok minősége és a közbeiktatott erősítőfokozatok miatt a digitalizálás után ez az érték 40–45 dB-re romlik.



Mivel vannak olyan rögzítő eszközök ebben a kategóriában, amelyek a törlést és egyben az előmágnesezést egy állandó mágnessel valósítják meg, ezért az előmágnesezés szintje közel sem lesz állandó, sokkal inkább jól hallható alacsony frekvenciájú előmágnesezési zaj lesz belőle. Ez a zajforma oly mértékben elnyújtott, és véletlenszerű, hogy algoritmikus csökkentése esetenként nem is lehetséges, ugyanis a zaj mintavételezése során nem választható ki úgy zajos rész, hogy e zajból kellő mennyiségű és minőségű jusson a mintába megfelelő kivonáshoz.

#### 4.2.2.3. Zajcsökkentési eljárás

A magnetofonszalagokról érkező jelek főként magas frekvenciájú zajt tartalmaznak, mely a korrekciós előerősítő fokozatból adódik. Ez a zaj mintavételezést követően spektrális kivonással csökkenthető a zavaró mérték alá.

Csak olyan esetekben ajánlott a magnetofonszalagok digitális hanghordozóra másolása, ha az anyag bír akkora fontossággal, hogy megtegyük (más hanghordozón, nem hozzáférhető). A magnetofonszalagokról tudjuk, hogy a normál szalag sávátvitele alig több mint a fele a krómdioxidos vagy magnetit szalagokéhoz képest, és a szalag zaja, ami a mágneses információhordozó minőségéből adódik, is lényegesen nagyobb az ilyen szalagoknál.

A magnetofonnál általában nem jelentkezik impulzív jellegű zaj, ezért a zajcsökkentés elvégezhető egyetlen lépésben. A digitalizált jelből (az elejéről vagy a végéről) zajmintát veszünk, majd pedig spektrális kivonással távolítjuk el a zajt. Az eredmény – bár az eljárás egy lépésből áll csupán – nem lesz olyan minőségű, mint a hanglemezek esetében, ugyanis a magasabb frekvenciájú hangokat a rögzítő rendszer és a szalag már korábban eltüntette, ezért lejátszáskor sem voltak már jelen. Ezek a hangok már semmilyen eszközzel nem állíthatók vissza.

## 5. Összegzés

Dolgozatom célja volt a környezetünkben előforduló zajok osztályozása, és azok csökkentésére megoldást találni. A potenciális zajforrások nagytöbbségére a dolgozatom első szakaszában világítottam rá. Tekintettel arra, hogy zaj hasonlóan az információt hordozó jelhez véletlenszerű, ezért modellezése és csökkentése közel sem olyan egyszerű, mint ahogyan azt a köznapi ember gondolja. A zajok modellezése a zajcsökkentés szempontjából jelentős, ugyanis egy megfelelő modell segítségével az eljárás könnyebben algoritmizálható, és alkalmazható a digitális jelfeldolgozásban. A számítógépek miniatürizálásával és gyors fejlődésével korunkban egyre nagyobb teret kap a kommunikációs és jelrögzítő rendszerekben a digitális jelek alkalmazása, mely a feldolgozhatóság és a hozzáadódó zajokkal szembeni jó „ellenálló-képességének” köszönheti sikerét.

A zajok modellezését kétféle szempont szerint közelítettem meg dolgozatomban, melyekből kiderült, hogy a véletlenszerű jelek előre meg nem jósolható értékeket is felvehetnek. A modellek alkalmazhatók más köznapi véletlenszerű események modellezésére is, melyek messze esnek az informatika tudományterületétől, mint például a közgazdaságtan, vagy a meteorológia, vagy az orvostudomány.

A probléma megoldása a zajcsökkentési eljárások alkalmazása. Többféle módszer létezik, melyekből csak néhányat volt alkalmam kiemelni, dolgozatom szűkös keretei miatt. A módszereket az adott problémának megfelelően, jelen esetben a zajok fajtáihoz illően kell alkalmazni, a várt eredmény elérése érdekében. A zajok csoportosítása, és forrásuk meghatározása ezért nagyon fontos, a legalkalmasabb eljárás kiválasztása szempontjából. Egy véletlenszerű jelsorozatból bizonyos esetben nehéz eldönteni még a szétválogatás után is, hogy melyik a számunkra hasznos információt hordozó jel. Erre nagyon szemléletes példa egy gépjármű motorja, mely az utas szempontjából zajforrás, de egy hozzáértő szerelő sok értékes információhoz juthat a hallott hangokból.

Az utolsó szakaszban az elméleti ismeretek gyakorlatba ültetését próbáltam kifejteni, saját tapasztalataim alapján. Több mint tíz évvel ezelőtt volt alkalmam több „házi stúdióban” készült könnyűzenei demófelvételhez segítséget nyújtani, és jómagam is aktív részese voltam egyes felvételek elkészítésének, mint hangszeres „zenész”. Jelenlegi iskolámban néha alkalmam nyílik kamatoztatni korábbi ismer-

reteimet, és dolgozatom megírásával nagyobb rálátást nyerhettem a későbbi munkámhoz a zajok forrásaira és megfelelő mértékű csökkentésükre.

A régebbi hanghordozókon tárolt, többnyire zenei felvételek korszerű adathordozóra mentése több ismerősömnek eszébe jutott már, de megfelelő eszköz hiányában maguk nem tudták elvégezni a műveletet. Érdeklődési körömbé tartozik a hangrendszerek felépítésének, és működésének megismerése és tanulmányozása, ezért több, mikrobarázdás lemez lejátszására alkalmas eszköz és magnetofon is található otthonunkban, mellyel a hanganyagok digitalizálásához forrást tudok biztosítani. A számítógépben található hangkártyákra mindig figyeltem, hogy ne csak gyenge minőségű hangjelet tudjon produkálni a komolyabb hangrendszer számára, hanem megfelelő működésével minőségi zenehallgatásra legyen alkalmas.

A dolgozat megírása során törekedtem arra, hogy a címben megragadott témát minél jobban, és több szempont szerint is kifejthessem, amit végeredményében sikerült megvalósítanom.

## 6. Irodalomjegyzék

1. Saaed V. Vaseghi – Advanced Digital Signal Processing and Noise Reduction  
Second Edition  
Copyright © 2000 John Wiley & Sons Ltd.
2. Valószínűségszámítás és statisztika előadások anyaga  
Nagy Márta – 2001–2002 tanév
3. Molnár-Sáska Gábor – Rejtett Markov modellek statisztikai vizsgálata  
Tézisfüzet – 2005  
forrás: [http://doktori.math.bme.hu/Molnar-Saska\\_tezisfuzet.pdf](http://doktori.math.bme.hu/Molnar-Saska_tezisfuzet.pdf)
4. dr. Tolnai János – Rádióamatőr vizsgára felkészítő tananyag  
Puskás Tivadar Távközlési Technikum, 2006.  
forrás: [http://www.puskas.hu/r\\_tanfolyam/r\\_tananyag.html](http://www.puskas.hu/r_tanfolyam/r_tananyag.html)

Debreceni Egyetem  
Informatika Kar

# Szakedolgozat

## Zajok és eltávolításuk

**Témavezető:** Dr. Papp Zoltán  
egyetemi adjunktus

**Készítette:** Gambár Péter István  
III. informatika tanár szakos,  
levelező tagozatos hallgató

Debrecen  
2009.

# Tartalomjegyzék

1. Bevezetés.....	4
1.1. A cím magyarázata.....	4
1.2. Témaválasztás indoklása .....	4
2. Zajok és modellezésük .....	5
2.1. Zajok fajtái és forrásai.....	5
2.1.1. Általános felosztás .....	5
2.1.2. Fehér zaj.....	7
2.1.3. Színezett zaj .....	7
2.1.4. Impulzív zaj.....	8
2.1.5. Tranziens zaj impulzus.....	10
2.1.6. Termikus zaj .....	11
2.1.7. Schottky-zaj.....	12
2.1.8. Elektromágneses zaj .....	13
2.1.9. Csatornatorzítások.....	14
2.1.10. Zajmodellezés .....	15
2.1.10.1. Hozzáadódó fehér normális eloszlású zajmodell (AWGN).....	16
2.1.10.2. Rejtett Markov modell.....	16
2.2. Zajmodellek .....	18
2.2.1. Valószínűségi modellek.....	18
2.2.1.1. Véletlenszerű jelek és sztochasztikus folyamatok.....	18
2.2.1.2. Valószínűségi modellek .....	19
2.2.1.3. Állandó és nem állandó véletlen folyamatok .....	23
2.2.1.4. Véletlen folyamatok várható értéke.....	24
2.2.2. A rejtett Markov modell.....	29
2.2.2.1. Nem stacionárius folyamatok statisztikai modellje .....	29
2.2.2.2. Rejtett Markov modellek.....	31
2.2.2.3. A rejtett Markov modell betanítása .....	34
3. Zajsűrítési módszerek.....	36
3.1. Zajzár .....	36
3.2. Analóg valós idejű zajsűrítők.....	36
3.2.1. Dolby zajsűrítő.....	36
3.2.2. Zajérvénytelenítés.....	38
3.3. Digitális jelek feldolgozása.....	39

---

3.3.1. Átalakítások .....	39
3.3.1.1. Analóg–digitális átalakítás .....	39
3.3.1.2. Digitális–analóg átalakítás .....	40
3.3.2. A Wiener szűrő.....	42
3.3.2.1. A legkisebb négyzetek becslése .....	42
3.3.3. Zajcsökkentés spektrális kivonással .....	44
3.3.3.1. A spektrális kivonás módszere.....	45
3.3.3.2. A spektrális kivonó megvalósítása.....	47
3.3.4. Az impulzív zaj eltávolítása.....	50
3.3.4.1. Impulzív zaj matematikai modellje .....	50
3.3.4.2. A medián szűrő.....	51
3.3.4.3. Impulzív zaj eltávolítása lineáris prediktív modell segítségével ...	53
4. Zajszűrés a gyakorlatban.....	58
4.1. Zenekari stúdiómunka .....	58
4.1.1. A felvétel .....	58
4.1.2. Az utómunka.....	59
4.2. Hagyományos hanglemez, hangkazetta archiválása CD-re.....	61
4.2.1. Hanglemez archiválása.....	61
4.2.1.1. A lejátszó rendszer .....	61
4.2.1.2. Várható zajok .....	63
4.2.1.3. Zajcsökkentési eljárás .....	64
4.2.2. Hangkazetta archiválása.....	65
4.2.2.1. A magnetofon és a hangszalag.....	65
4.2.2.2. Várható zajok .....	66
4.2.2.3. Zajcsökkentési eljárás .....	67
5. Összegzés .....	68
6. Irodalomjegyzék.....	70

## 1. Bevezetés

### *1.1. A cím magyarázata*

Általában hangfelvételek (függetlenül attól, hogy milyen eszközzel készülnek és milyen adathordozóra) és kommunikációs rendszerek esetében felütik fejüket az additív zajok. Ezen zajokat a hasznos jelhez szokták viszonyítani, és sajnos adott esetben a hasznos jel nagyságát is meghaladhatják, mely nagyban károsítja a jelek értelmezhetőségét. Több esetben azonban lehetőség nyílik az ilyen zajok mértékének, erősségének csökkentésére, mind hagyományos, mind digitális rendszerben. A zajok eltávolításánál nagyon fontos, hogy a hasznos jel minél nagyobb hányada maradjon ép, értelmezhető, és lehetőleg torzításmentes. Zajforrástól függően a zajok csökkentése hagyományos úgynevezett analóg rendszerekben is megvalósítható, akár megfelelő alkatrészek és eszközök alkalmazásával is. Jelen dolgozatban a tárolt és szállított hangjelek zajainak felismerésével, fajtaival és azok csökkentésével foglalkozom.

### *1.2. Témaválasztás indoklása*

Több éve egy zenei általános iskolában tanítok, ahol esetenként szükség van a kórus munkájának megörökítésére nem csak képi, de hang formájában is. A hangfelvételek megfelelő minőségű elkészítése és tárolása érdekében a rögzítő technika és az előtte található átalakító rendszerek által termelt zajok eltüntetése adott esetben nagyon fontos. Célom, hogy a dolgozat megírásával iskolámban minél hatékonyabb munkát tudjak végezni ebben a tekintetben is.

Személyes érdeklődési körömbé tartozik a zenei hangfelvételek digitális feldolgozása, illetve az erre alkalmas berendezések és szoftverek megismerése is, tehát többek között ez is motivált a szakdolgozat témájának megválasztásánál.



## 2. Zajok és modellezésük

### 2.1. Zajok fajtái és forrásai

#### 2.1.1. Általános felosztás

A zaj definiálható olyan nem kívánt jelnek, mely megakadályozza (zavarja) egy másik jellel (hasznos jel) történő kommunikációt vagy a másik jel mérését. A zaj maga is egy információhordozó, a zajforrás szempontjából. Például egy autó motorjának zaja információval szolgál a motor állapotáról. A zajok igen sokfélék és változatosak lehetnek, kezdve a hangfrekvenciás akusztikus zajoktól a rádiófrekvenciás elektromágneses zajokig.

A másik említésre méltó probléma a hasznos jel kiindulási alakjától való nem kívánatos eltérése, torzulása. Ebbe a csoportba tartoznak a különféle visszhangok, felharmonikusok, és az elveszett jelerészetek.

A zajok és torzulások a kommunikációs és mérési rendszerek fő ellenségei, ezért e káros mellékhatások eltüntetése érdemel kiemelt figyelmet. A zajszűrést és torzítás kiküszöbölését főként a mobil kommunikációban, a beszéd felismerésben, orvosi jelfeldolgozásban, radar és hanglokátor rendszerekben alkalmazzák a köznap életben.

A zajok életünk minden területén, minden környezetben jelen vannak, különféle formában. Az zaj forrásától függően a következő nagy csoportok különíthetők el egymástól:

- a.) *akusztikus zaj*: leggyakrabban mozgásból, vibrációból (rezgésből), ütközésből származik, ezáltal ez a zajforma – eltérő mértékben ugyan – jelen van a mindennapi környezetünkben. Ezt a zajt legtöbbször haladó gépjárművek, légkondicionálók, számítógépek hűtőventillátorai, közlekedés, emberi beszélgetések, szél, eső, stb. okozzák.
- b.) *elektromágneses zaj*: minden frekvencián jelen van, különösen a rádiófrekvenciákon. Minden elektromos berendezés, például a rádió- és TV adóállomások, és vevőkészülékek állítanak elő jelentős elektromágneses zajt.
- c.) *elektrosztatikus zaj*: feszültség jelenléte okozza, áram nélkül vagy folyó árammal. Ennek a zajnak az egyik legjelentősebb forrása a fénycsővilágítás.

d.) *csatornatorzítás, visszhang, és csillapítás*: az átviteli csatorna nem ideális karakterisztikája következtében történik. A rádiófrekvenciás csatornák, melyeket a mobil kommunikációban használnak, – mikrohullámú frekvenciákon – különösen érzékenyek a terjedés karakterisztikájára és környezetére.

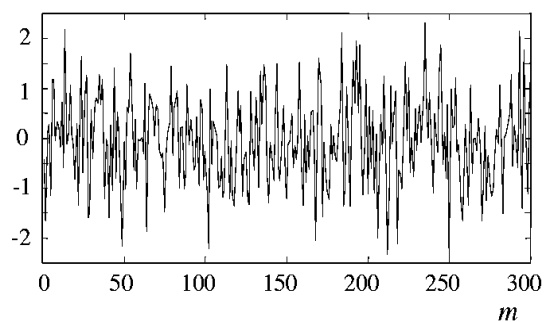
e.) *feldolgozási zaj*: az a zaj, amelyet a jel digitális–analog átalakítása eredményez, például a zaj digitálisan kódolt beszédben vagy hasonló jelekben, vagy az elveszett csomagok a digitális adatkommunikációs rendszerekben.

A zaj frekvenciájának vagy időkarakterisztikájának függvényében a következő kategóriákba csoportosítható:

- *Keskeny hullámsávú zaj*: olyan keskeny hullámsávon érkező zajhatás, mint az 50 Hz-es vagy 60 Hz-es brummogás (búgás) mely többnyire az elektromos tápegységekből, váltakozó áramú vezetékekből származik.
- *Fehér zaj*: tisztán véletlen zaj, melynek teljesen egyenes az energiaspektruma. A fehér zaj elméletileg az összes frekvenciát tartalmazza, azonos intenzitással.
- *Sávkorlátozott fehér zaj*: egyenes spektrumú zaj korlátozott sáv szélességgel, mely általában az alkalmazott eszköz korlátozott szélességű sávját, vagy az általunk használt sávot jelenti.
- *Szinezett zaj*: nemfehér zaj, vagy egyéb szélessávú zaj, melynek nem egyenes a spektruma, például: rózsaszín zaj, barna zaj, autoregresszív zaj.
- *Impulzusszerű zaj*: véletlenszerű rövid időtartamú impulzusokból áll, véletlenszerű amplitúdóval, hosszal és felbukkanással.
- *Tranziens zaj*: viszonylag hosszú időtartamú impulzusokból áll.

### 2.1.2. Fehér zaj

A fehér zaj definíció szerint egy összefüggéstelen jelekből álló zaj, mely minden frekvencián azonos energiájú. A zaj, melynek energiája minden frekvencián azonos  $\pm \infty$  tartományban, szükségképpen végtelen nagy energiájúnak kellene lennie, ennek következtében ez csak egy elméleti fogalom. Habár egy sávkorlátozott zaj,



1. ábra Fehér zaj

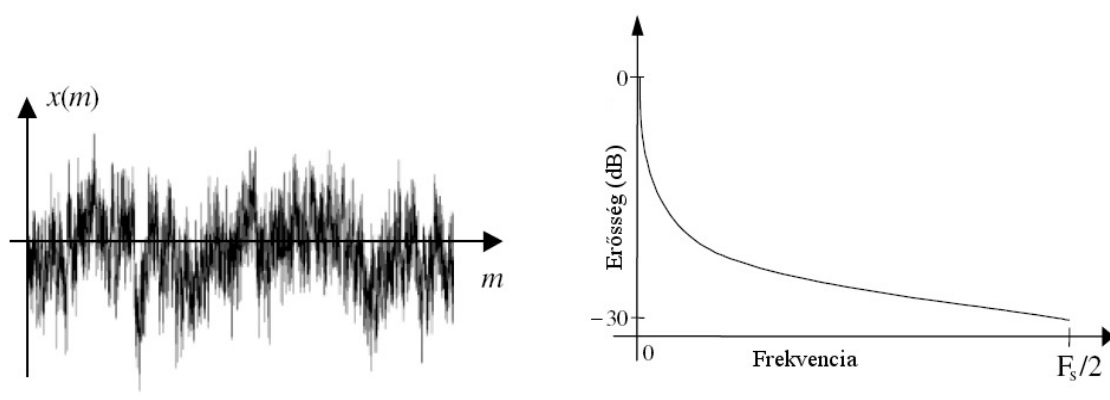
egyenes spektrummal – ha a zaj a teljes vizsgált sávot lefedi, amit a kommunikációs rendszer biztosít – a rendszer szempontjából minden tekintetben fehér zajnak számít. Például egy 10 kHz sávátvítellel rendelkező hangrendszer számára minden 10 kHz-nél szélesebb sávú egyenes spektrumú hallható zaj fehér zajnak

tűnik.

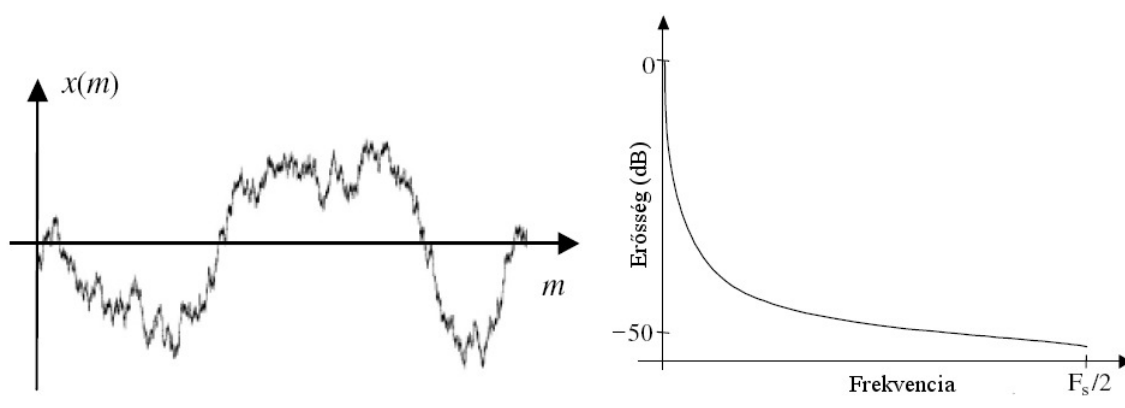
A tiszta fehér zaj egy elméleti fogalom, mivel végtelen energiájúnak kellene lennie a végtelen frekvenciatartomány lefedésére. Ráadásul a diszkrét időpontokban mintavételezett jelnek szükségképpen sávkorlátozottnak kell lennie, ahol a legnagyobb frekvenciaérték kisebb, mint a mintavételezési frekvencia fele. A sávkorlátozott fehér zaj egy sokkal gyakorlatibb fogalom, mely definíció szerint egy egyenes spektrumú zaj egy adott frekvenciasávban. Egy ilyen zaj energiája nem végtelen, ezért kiszámítható.

### 2.1.3. Színezett zaj

Bár a fehér zaj elmélete meglehetősen valós, és matematikailag kényelmes és használható közelítést ad több zajjal kapcsolatban, melyek felütik fejüket a telekommunikációs rendszerekben, sok egyéb zaj viszont nem fehér. A színezett zaj kifejezés bármely olyan szélessávú zajra utal, melynek spektruma nem a fehér zajéval egyező. Például a legtöbb hangfrekvenciás zaj, mint a mozgó autók zaja, a számítógépek ventilátorainak zaja, a fűrés zaja és emberi beszélgetés zaja, nemfehér, és túlnyomórészt alacsonyfrekvenciájú hangokat tartalmaz. Egy fehér zaj adott csatornán történő továbbításakor a zaj bizonyos frekvenciákon történő csillapodásával jellemezhető a „színezett zaj”, a csatorna spektrumának alakjától függően. A színezett zaj két klasszikus változata a rózsaszín (2. ábra) és a barna (3. ábra) zaj.



2. ábra Rózsaszín zaj és spektruma



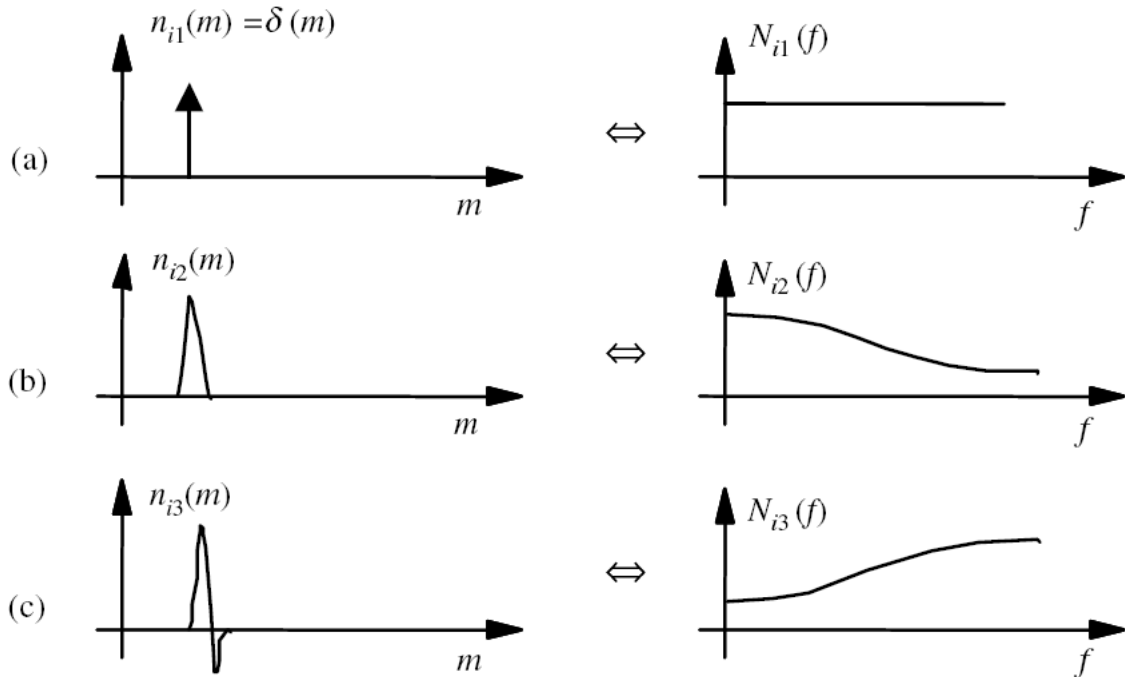
3. ábra Barna zaj és spektruma

#### 2.1.4. Impulzív zaj

Az impulzív zaj rövid időtartamú „be/ki” impulzusokból áll. Ezek a zajok többféle forrásból származhatnak, mint például kapcsolási zaj, kedvezőtlen csatorna-környezet, kimaradás vagy felületi kopás a hanghordozón, számítógép-billentyűzet kattogása, és így tovább. A 4. ábra (a) része egy ideális impulzív zajt mutat, annak spektrumával. A kommunikációs rendszerekben előforduló impulzív zaj jellemző paramétere a hossza, mely normális esetben több mint egy minta hosszúságú. Például hangrendszerek esetében a rövid időtartamú, éles impulzusok lehetnek akár 3 milliszekundum (60 minta 20 kHz-es mintavételezési frekvencia esetében) hosszúak, melyeket impulzív zajnak tekinthetünk. A 4. ábra (b) és (c) része két példát mutat rövid időtartamú impulzusokra és azok spektrumára.

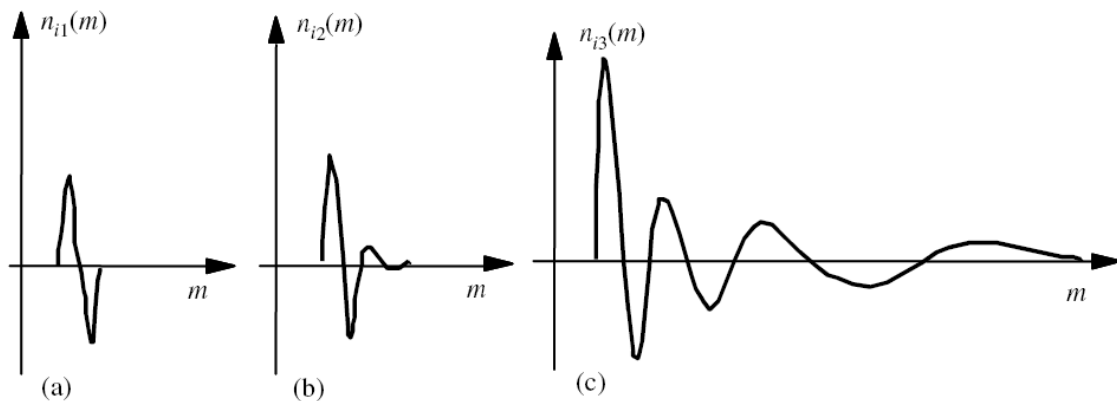
Egy kommunikációs rendszerben az impulzív zaj a térben és időben több pontból ered, azután a csatornán keresztül halad a vevő felé. A vett zaj időben szétszórta és a csatorna által alakított, és úgy kezelhető, mint a csatorna impulzus jelre adott reakciója. Általában a csatornák karakterisztikája lehet lineáris vagy nem lineáris,

időben állandó, vagy időben változó. Azonkívül sok kommunikációs rendszer válaszul a nagy amplitúdójú impulzusokra nem lineáris karakterisztikát mutat.



4. ábra Impulzív zaj

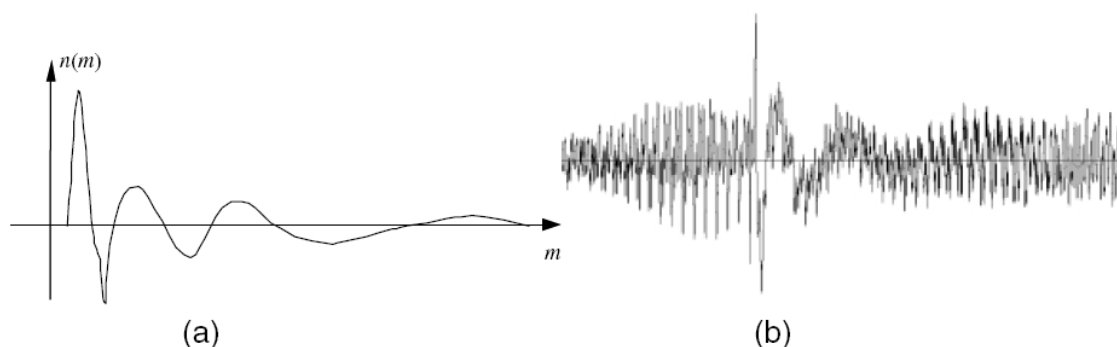
Az 5. ábra néhány példát mutat impulzív zajra, melyek tipikusan régi hanglemezeken fordulnak elő. Ebben az esetben a kommunikációs csatornának a lejátszó rendszert tekintjük, melyet időben állandónak feltételezhetünk. Az ábra három változatát mutatja a zaj-amplitúdóváltozásának az idő függvényében. Az 5. ábra (c) eleme mutatja, hogy egy nagy impulzslökés csillapodó tranziens impulzussorozatot hoz létre. Ezek a példák a lejátszó rendszer nem lineáris karakterisztikáját jelzik.



5. ábra Példák impulzív zajra

### 2.1.5. Tranziens zaj impulzus

Az tranziens zaj gyakran áll egy viszonylag rövid éles impulzusból, melyet egy csillapodó viszonylag alacsony frekvenciájú oszcilláció követ, ahogy azt a 6. ábrán is láthatjuk. A kezdő impulzust gyakran valamilyen belső vagy külső interferencia okozza, míg az oszcillációt gyakran a csatorna rezonanciája, melyet a kezdeti impulzus gerjeszt, és tekinthető a csatorna válaszának a kezdeti impulzusra. Egy telekommunikációs rendszerben az impulzív zaj a térben és időben több pontból származik, azután a csatornán keresztül halad a vevő felé. A vett zaj időben szétszórt és a csatorna által torzított, alakított és úgy tekintendő, mint a csatorna egy impulzusra adott válaszreakciója. Az impulzív zaj a kezdő impulzussal és az azt követő oszcillációval jellemezhető, a csatorna szempontjából.



6. ábra Egy karc impulzus és zene egy hanglemezről (b), hanglemez karc impulzusának átlagolt profilja (a)

A 6. ábra jól szemlélteti a tranziens zaj impulzus alakját, tekintsük meg a karcolás impulzust<sup>1</sup> az ábrán, mely egy sérült hanglemezről származik. A karcolás impulzusok a lemeztű és a hozzá kapcsolt elektromechanikus lejátszó rendszer akusztikus válaszai a lemezen található fizikai megszakítások jelenlétére. Mivel a karcolás hatására előálló tranziens zajok tulajdonképpen a lejátszó rendszer impulzus válaszai, elvárható, hogy az adott rendszer különböző helyeken található impulzusokra hasonló karakterisztikát mutasson. Ahogyan a 6. (a) ábra is mutatja egy tipikus karc impulzus hullámalak gyakran két különböző részre tagolódik:

- egy bevezető nagy amplitúdójú lökésre, mely a lejátszó rendszer válasza a rögzített médiumon (lemezen) található karcolásra, melyet
- csillapodó hullámzások követnek, további torzítást okozva ezzel.

<sup>1</sup> Karc(olás) impulzus alatt egy hanglemezen, a használat során kialakult (többnyire a sávval valamilyen szöveget bezáró) karc (tekinthető pillanatnyi jelfolytonosság-hiánynak is) hatására létrejövő lejátszáskor jelentkező ugrásszerű lökést érthetünk, mely közvetlenül a mechanikus jelérzékelőre hat, és a mechanikai–elektronikai átalakítást követően végighalad a teljes elektronikus rendszeren is.

A kezdő impulzus általában viszonylag rövid, és a hossza a 1–5 ms tartományba esik, míg a hullámzó farok hosszabb időtartama elérheti akár az 50 ms-ot vagy többet is.

A 6. (b) ábrán látható csillapodó hullám (farok) frekvenciája is csökken az idő múlásával. Ez a viselkedés az elektromechanikus lejátszó-rendszer nem lineáris válaszmechanizmusa a lemezen található karcokra. Tapasztalat alapján több hanglemezről származó recsenés zaja jól meghatározható hullámalakot mutat, és viszonylag kevés tipikus sablonnal jellemezhető.

### 2.1.6. Termikus zaj

A termikus zaj – más néven Johnson-zaj (felfedezőjéről J. B. Johnsonról<sup>1</sup> elnevezve) – hővel stimulált részecskék rendezetlen mozgásából adódik. A termikus zaj elméletének alapjai a termodinamikában keresendők. A termodinamikában a szabad részecskéket hőmérsékletfüggő rendezetlen mozgás jellemzi, mint például a szabad gázmolekulák egy tartályban, vagy a szabad elektronok egy vezetőben. Bár ezen rendezetlen mozgások átlagértéke közel nulla, az átlag körüli hullámzások alkotják a termikus zajt. Például a gázmolekulák rendezetlen mozgásai és ütközései egy behatárolt térben rendezetlen hullámzást okoznak az átlagos nyomás körül. A hőmérséklet növekedésével a molekulák vagy más elemi részecskék kinetikai energiája növekszik, és ezzel a termikus zaj is.

Ehhez hasonlóan egy elektromos vezetőben nagy számú szabad elektron van jelen (az ionokkal együtt), melyek rendezetlenül rezegnek az egyensúlyi pozíciójuk körül, és útját állják a további elektronok mozgásának. Az elektronok rendezetlen szabad mozgásai spontán áramokat, vagy más szóval termikus zajt képeznek, melynek átlaga nulla, mivel feszültség hiányában minden elektron mozgása különböző irányú (nincs egyirányú áramlás). A környezet hőmérsékletének növekedésével a vezető hőmérséklete is növekszik, mely az elektronokat magasabb energiaszintre viszi, ezzel növelve a rendezetlen áramfolyásokat. Egy fémréteg ellenállásban a pillanatnyi feszültségek négyzetének átlagértéke a termikus zaj miatt a következő összefüggéssel adható meg:  $\overline{v^2} = 4kTRB$ , ahol  $k = 1,38 \cdot 10^{-23} \frac{\text{J}}{\text{°K}}$  a Boltzmann-állandó,  $T$  az abszolút hőmérséklet °K-ben,  $R$  az ellenállás ohmban, és  $B$  a sávszélesség. Az iménti egyenletből és az azt megelőző érvelésből egy áramkörü

<sup>1</sup> John Bertrand “Bert” Johnson (1887–1970): svéd születésű amerikai villamosmérnök és fizikus. Ő fejtette ki először részletesen a kábeleken közvetített információ véletlen interferenciájának alapjait. (Wikipédia – [http://en.wikipedia.org/wiki/John\\_B.\\_Johnson](http://en.wikipedia.org/wiki/John_B._Johnson))

lapba ültetett fémréteg ellenállás úgy viselkedik, mint egy termikus zajgenerátor, mely feszültségnégyzetének átlaga  $\overline{v^2}$ , és a belső ellenállása  $R$ . Az ily módon előálló termikus zajgenerátor teljesítménye számítható ki a következő összefüggéssel:

$$P_N = \overline{i^2} R = \left( \frac{v_{rms}}{2R} \right)^2 R = \frac{\overline{v^2}}{4R} = kTB[\text{W}], \text{ ahol } v_{rms}^1 \text{ az effektív feszültség értéke. A ter-}$$

mikus zaj spektrumsűrűsége megadható a  $P_N(f) = \frac{kT}{2} \left[ \frac{\text{W}}{\text{Hz}} \right]$  összefüggéssel. Ebből

az egyenletből kiderül, hogy a termikus zaj spektrumsűrűsége egyenes, azaz a termikus zaj egy fehér zajnak tekinthető (mely általában sávkorlátozott). Ezt az összefüggést különösen a magas rádiófrekvenciák ( $10^{13}$  Hz) esetében fontos.

### 2.1.7. Schottky-zaj

A Schottky-zaj<sup>2</sup> elmélete az elektroncsövek katódjának elektronkibocsátás-ingadozásából származik. Az különállónak tekintett elektronok áramlás során véletlenszerű időpillanatokban érkeznek, ezért az átlagérték körüli ingadozást okoznak a részecskefolyamban. Ezek az ingadozások a részecskefolyam sebességében Schottky-zajt okoznak. Egyéb példák Schottky-zajra a fotonok áramlása egy lézer fénysugárban, elektronok és lyukak áramlása valamint rekombinációja<sup>3</sup> félvezetőben, vagy a fotóelektronok árama, melyek a fotódiódákban emittálódnak. A véletlenszerűség jelenléte a részecske-kibocsátás vagy -érkezés mértékében magába foglalja, hogy a Schottky-zaj Poisson eloszlással modellezhető. A hosszú megfigyelési idejű kísérletek esetén az érkezések átlagérték körüli ingadozása a normális eloszláshoz közelít. Míg a termikus zaj a részecskék „nyugalmi állapotban” történő véletlenszerű mozgásából adódik, addig a Schottky-zaj egy (feszültséggel) irányított részecske-áramlásra értelmezhető.

<sup>1</sup>  $v_{rms}$  – root mean squared voltage: a feszültség effektív értéke

<sup>2</sup> Walter Hermann Schottky német fizikus, a film-rácsú vákuumcső (1915) és a tetróda (1919) feltalálója – munkáit a Siemensnél végezte. 1938-ban megfogalmazott egy elméletet, mely a Schottky-hatást jelezte, ma pedig a Schottky-diódákban használjuk. – [http://en.wikipedia.org/wiki/Walter\\_H.\\_Schottky](http://en.wikipedia.org/wiki/Walter_H._Schottky)

<sup>3</sup> „A félvezető nem csak szabad elektronjai útján vezeti az áramot (mint a fémek), hanem egy másik vezetési mechanizmus: a lyukvezetés útján is. Az eredetileg elektromosan semleges atom a negatív töltésű elektron távozása folytán pozitív töltésűvé válik, ezért a lyukat pozitív töltésűnek tekintjük. Előfordulhat, hogy egy szomszédos atom valamelyik vegyértékelektronja szintén elszabadul, és „beugrik” az előbb említett (pozitív töltésénél fogva az elektront vonzó) lyukba. Ezzel az előbbi lyuk (és az oda beugró szabad elektron is) megszűnik, hiszen a rácsszerkezet rendje az adott helyen helyreállt (ezt a folyamatot nevezik *rekombinációnak*), viszont a most elszabadult szabad elektron helyén új lyuk keletkezett. Ha valamelyik szomszédos atomból elszabaduló elektron most ebbe a lyukba ugrik be, tehát rekombinálódik, a lyuk a most elszabadult atom helyére »vándorol át.«” Forrás: Internet – [http://www.puskas.hu/r\\_tanfolyam/felvezetok.pdf](http://www.puskas.hu/r_tanfolyam/felvezetok.pdf), 2. oldal



Tekintsünk egy elektromos áramot, mint különálló elektromos töltéshordozók áramlását. Ha a töltéshordozók egymástól függetlenül mozognak, akkor az áram ingadozása megadható a  $I_{Noise}(rms) = \sqrt{2eI_{dc}B}$  összefüggéssel, ahol  $e = 1,6 \cdot 10^{-19}$  coulomb (egy elektron töltése), és  $B$  pedig a mérhető sávszélesség. Például egy  $I_{dc}$  1 amper „állandó” áram, és 1 MHz sávszélesség mellett, az effektív hullámzás 0,57 mikroamper lesz. Az iménti összefüggés feltételezi, hogy egymástól független töltéshordozók alkotják az áramot. Abban az eset mikor a töltéshordozóknak egy akadályon kell áthaladniuk, mint például egy félvezető dióda p–n átmenete, a töltéshordozók nem együtt mozognak; viszont nem igaz ez a fém vezetőkre, ahol a töltéshordozókat nagyfokú kölcsönhatás jellemzi.

### 2.1.8. Elektromágneses zaj

Virtuálisan minden elektromos eszköz, mely energiát állít elő, energiát fogyaszt vagy energiát szállít, elektromágneses zajok potenciális forrása. Az ilyen elektromágneses zaj interferenciának tekinthető más rendszerek szempontjából. Általában minél nagyobb a feszültség vagy az áram és minél közelebb helyezkedik el az áramkörhöz vagy elektromos eszközhöz, annál nagyobb lesz az áramkörben tapasztalható elektromágneses zaj mértéke. Az elektromágneses zajok általános forrásai a transzformátorok, rádió és televízió adóállomások, mobiltelefonok, mikrohullámú adóállomások, váltakozó áramú vezetékek, motorok és motorindítók, generátorok, relék, oszcillátorok, fénycsövek, és villámlások (zivatarok).

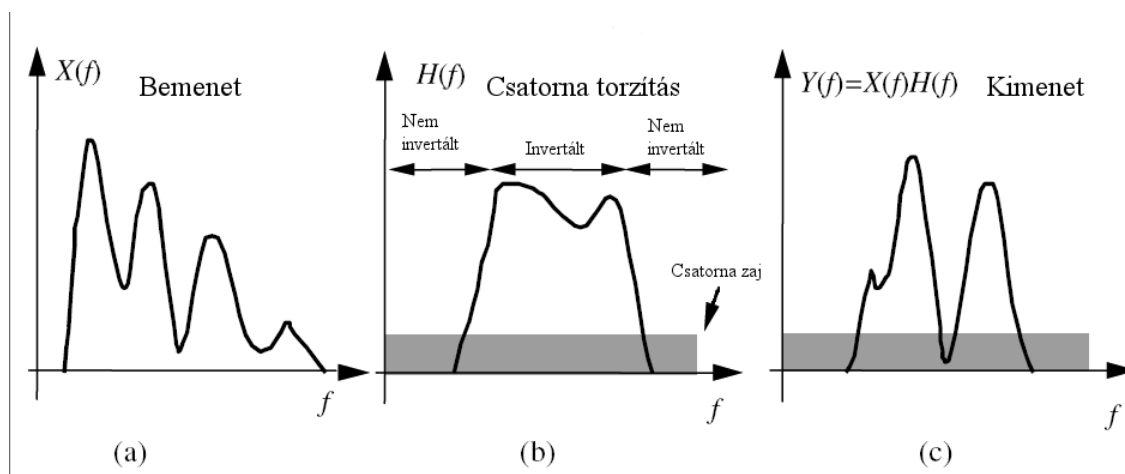
Az elektromágneses zajok két alapvető csoportba sorolhatók: elektrosztatikus zajok, és mágneses zajok. Ez két típus alapvetően különbözik egymástól, ezért más-más zajvédelem szükséges hozzájuk. Sajnos az imént felsorolt zajforrások legtöbbje a két típus kombinációját állítja elő, mely tovább bonyolítja a zajcsökkentési problémát.

Az elektrosztatikus tereket a feszültség jelenléte generálja, árammal, vagy anélkül. Az elektrosztatikus zajok egyik legáltalánosabb forrásai a fénycsövek. Mágneses terek keletkeznek elektromos áram hatására, vagy állandó mágnes jelenléte esetén. A motorok és a transzformátorok szolgálhatnak például az iménti felsorolásból, és hogy ne menjünk messzire a Föld mágneses tere is itt van környezetünkben. Ahhoz, hogy zajfeszültség indukálódjon egy vezetőben, a mágneses erővonalakat metszenie kell. A generátorok működési elve is ezen a jelenségen alapul. A váltakozó mágneses tér jelenléte – amilyen például az 50/60 Hz-es elektromos vezetékeket körülveszi – feszültséget indukál minden álló vezetőben, ahogy a mág-

neses tér kialakul, majd eltűnik (váltakozó mágneses tér). Ehhez hasonlóan, ha egy vezető mozog a Föld mágneses terében, zajfeszültség indukálódik benne, ahogy mozgás közben metszi a Föld mágneses erővonalait.

### 2.1.9. Csatornatorzítások

Ha egy jelet egy csatornán keresztül továbbítunk, a jel alakja változik, torzul a csatorna frekvenciaátvivő képességének és csillapítási paramétereinek megfelelően. A csatornatorzítás két fő megnyilvánulási formája a folytonossági torzítás és a fázistorzítás. A rádiókommunikációban létezik egy multi-út hatás is, mely szerint ha a jel több átjátszón megy keresztül, bizonyos esetekben különböző utakon haladva, ezzel a jel többféle változata érkezik a vevőbe különböző késleltetéssel és csillapítással. A csatornatorzítások csökkenthetők, vagy egészen egyszerűen szétválaszthatók a kommunikációs folyamatokban, ezért a csatorna modellezése és ki-

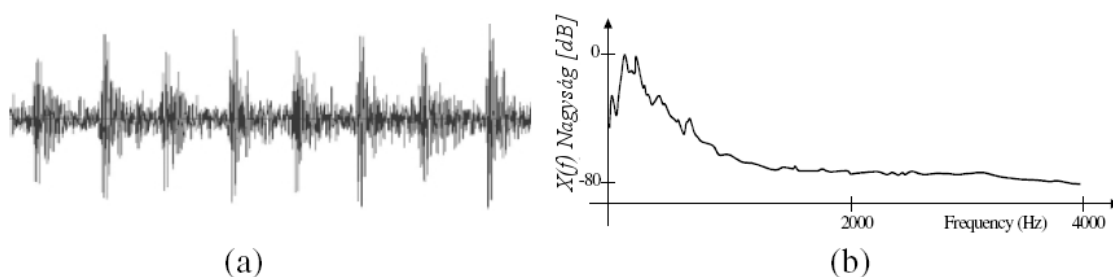


7. ábra Csatornatorzítás

egyenlítése nagyon jelentős összetevői egy modern digitális kommunikációs rendszer kialakításának. A csatorna kiegyenlítése különösen fontos a modern sejtes felépítésű kommunikációs rendszerekben mivel a csatorna karakterisztikák és terjedési paramétereik különbsége igen változatos, sokkal változatosabb, mint a vezetékes rendszerekben. A 7. ábra egy csatorna sávátviteli jellemzőit mutatja, mely két nem megfordítható és egy megfordítható régióra oszlik. A megfordítható csatornarégióban a jelek erősen torzultak elvesztek a csatornazaj miatt. A nem megfordítható csatornarégióban a jelek torzultak ugyan, visszanyerhetők. Ez a példa remekül mutatja, hogy a csatornafordító szűrőket nagy gondossággal kell elkészíteni, a nem kívánt eredmények elkerülése érdekében, mint például a zajerősítés az alacsony jel-zaj arányú frekvenciákon.

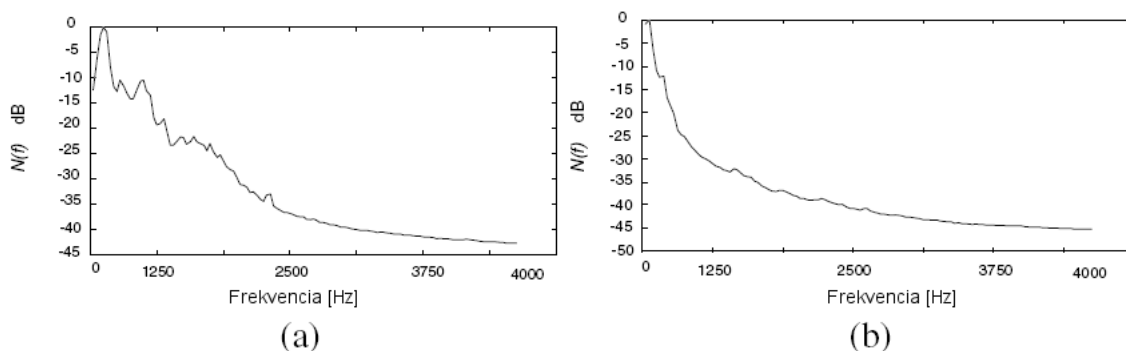
## 2.1.10. Zajmodellezés

A cél a modellezés esetén, hogy a jelben vagy zajban jelalakokat és mintákat jellemezzünk. A zaj pontos modellezéséhez egy olyan struktúrára van szükségünk, mely mind időben mind spektrumában jellemzi a zajt. A pontos zajstatisztika modellezéshez a kulcs a zajos jel megfelelő osztályozása és a hasznos jel erősítése. Ezen látszólag egyszerű jel/zaj osztályozási munka döntően függ a rendelkezésre álló hasznos jel mennyiségétől, erősségétől és a zajmodellektől, és ezen modellek alkalmazásától.



9. ábra A fúrózaj (a) idő-hullámformája (b) frekvenciaspektruma

A leghasználhatóbb szinte nélkülözhetetlen módszer az összes közül a zaj felépítésében való éleslátásra a Fourier transzformáció a zaj frekvenciaanalizálására. A 8. ábra egy elektromos fúró zaját mutatja, és ahogy az várható a zaj periodikus. Az ábra (b) része mutatja a fúró zajának spektrumát, melyen látható, hogy a zaj energiájának jelentős része a spektrum alacsonyabb frekvenciáira koncentrálódik. Tényszerűen igaz, hogy a legtöbb hangfrekvenciás jelnek és zajnak van túlnyomórészt alacsonyfrekvenciás a spektruma. Bár meg kell jegyezni, hogy a viszonylag kisebb energiájú magasabb frekvenciájú hangjelek fontos szerepet játszanak az élmény és a minőség átadásában. A 9. ábra (a) és (b) része példával szolgál az autók belsejéből származó zaj spektrumából (egy BMW és egy Volvo zajai láthatók). A



8. ábra Autók zajának spektuma (a) BMW 112 km/h, (b) Volvo 112 km/h

zajok nem statikusak, változatosak, és a következő forrásokból származhatnak:

- a.) kvázi periodikus zaj a motorból, és a jármű forgó alkatrészeiből,
- b.) zaj a kerekek és a talaj érintkezéséből és az út felületéből adódóan,
- c.) az autó körüli légáramlatok a légjáratokon, ablakokon, napfénytetőn keresztülhaladva,
- d.) az éppen előzőtt jármű zaja

Az autó zajainak jellemzői változnak a sebességgel, az út felületének állapotváltozásával, az időjárással és az autó környezetével.

A zajmodellezés legegyszerűbb módja, melyet gyakran használunk a jelenlegi gyakorlatban, hogy kiszámítjuk a zaj statisztikáját jelmentes időszakok alapján. Optimális esetben a Bayes féle jelfeldolgozási módszerekben egy készlet valószínűségi modell van előkészítve a jel- és zajfeldolgozásra. Ezek a modellek alapjául szolgálnak a jelek és zajok dekódolásának, zajos jelek felismerésének, erősítésének.

#### 2.1.10.1. Hozzáadódó fehér normális eloszlású zajmodell (AWGN<sup>1</sup>)

A kommunikációelméletben a gyakran feltételezzük, hogy a zaj hozzáadódó fehér normális eloszlást követő. Bár számos probléma esetében ez valós feltevés, és matematikailag kényelmes és használható megoldáshoz vezet, a gyakorlatban a zaj gyakran időben változó, kölcsönösen összefüggő, és nem normális eloszlást követő. Ez részben igaz az impulzív típusú zajokra, melyek nem állandóak, nem normális eloszlásúak, ezért nem modellezhetők az AWGN feltevés használatával. A nem állandó és nem normális eloszlású zajok modellezésére a Markov-féle sorozat alkalmas, stacionárius alfolyamatok leírásával.

#### 2.1.10.2. Rejtett Markov modell

A legtöbb zaj nem állandó, ez érvényes a zaj statisztikai paramétereire, úgy mint a zaj középértéke, szórásnégyzete, erőssége és spektruma, és ezek időbeli változása. A nem stacionárius zajok rejtett Markov modellek (HMM<sup>2</sup>) segítségével modellezhetők. Egy HMM lényegében stacionárius alfolyamatok véges sok állapotú Markov sorozata. Az állandó zaj modellezhető egy egyállapotú HMM segítségével. A nem állandó zaj modellezésére a többállapotú HMM alkalmas, amely képes a zaj időbeli változásának követésére, véges sok állandó állapottal. Nem normális eloszlású zajok esetében egy Gauss sűrűségi modell keverékét alkalmazhatjuk, a zaj terének modellezésére minden állapotban. Általában az állapotok száma modellen-

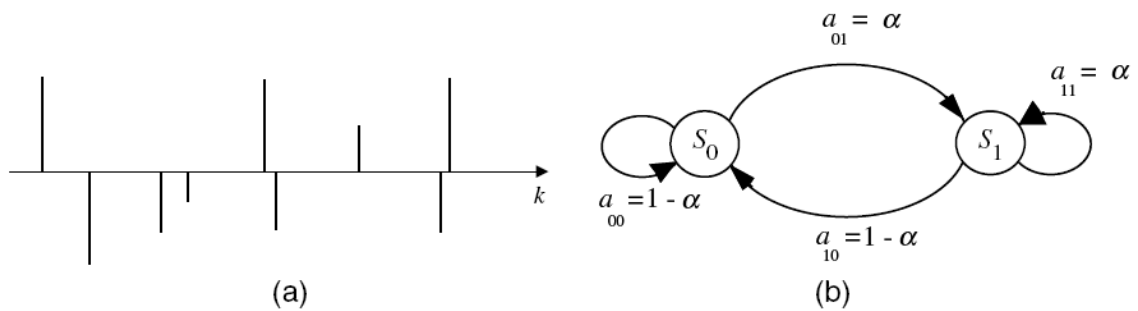
---

<sup>1</sup> AWGN: Additive White Gaussian Noise Model

<sup>2</sup> HMM: Hidden Markov Model

ként, és a keverékek száma állapotonként szükséges a zaj pontos modellezéséhez, mely a nem állandó zaj karakterisztikájától függ.

Egy példa a nem állandó zajra az impulzív zaj, melyet a 10. ábra (a) része szemléltet. Az ábra (b) részén az impulzív zaj szekvencia egy kétállapotú HMM modellje látható, ahol az  $S_0$  állapot jelképezi az „impulzus-ki” periódusokat az impulzusok között, és az  $S_1$  állapot jelképezi magát az impulzust. Azokban az esetekben ahol minden impulzus jól definiált átmeneti felépítéssel rendelkezik, előnyös lehet a többállapotú HMM magának az impulzusnak a modellezésére.



10. ábra (a) Impulzív zajszekvencia, (b) az impulzív zaj kétállapotú modellje

## 2.2. Zajmodellek

### 2.2.1. Valószínűségi modellek

A valószínűségi modellek alapja az információelmélet. Az információ maga is mennyiségileg a valószínűség logaritmusaként fejezhető ki. A valószínűségi modellek általában azzal foglalkoznak, hogy történések előfordulását megjósolják és jellemzésük véletlenszerű események alapján, különféle területeken, mint például telefonhívások számának megjósolása egy fővonalon a nap egy meghatározott időszakában, közúti forgalommodellezés, időjárás előrejelzés, pénzügyi adatmodellezés, előre jelezni a drogok hatását megadott adatok alapján az orvosi gyakorlatban. A jelfeldolgozásban a valószínűségi modelleket általában véletlenszerű jelek ingadozásának leírására alkalmazhatjuk, mint például mintafelismerés, jelkódolás és jelkiértékelés.

#### 2.2.1.1. Véletlenszerű jelek és sztochasztikus folyamatok

A jelek egyik legalapvetőbb jellemzőjük szempontjából két nagy csoportba osztályozhatók: determinisztikus (pontosan meghatározott) jelek és véletlenszerű jelek. Az időben véletlenszerű jelekre gyakran úgy hivatkozunk, hogy sztochasztikus jelek. Minden osztályban a jelek lehetnek folyamatosak és diszkrét az időben, és rendelkezhetnek folytonos értékű és diszkrét értékű amplitúdóval.

A determinisztikus jel definiálható úgy, hogy egyike az előre meghatározott görbéjű jeleknek időben és térben. A determinisztikus jelek pontos hullámszáma teljes mértékben leírható az idő függvényében, és a jel pontos értéke bármely időpillanatra kiszámítható a függvényi leírásból, és a korábbi értékek is kiszámíthatók ezen a módon. Például egy szinusz hullám  $x(t)$  modellezhető és pontosan kiszámítható mind egy másodrendű lineáris prediktív modellből, vagy a közismertebb  $x(t) = A \cdot \sin(2\pi ft + \phi)$  egyenlettel.

A véletlenszerű jeleknek nem kiszámítható hullámszámai vannak, ezért nem lehet olyan egyenletet felírni, mellyel kiszámítható lenne a jel pontos jövőbeli értéke a korábbi értékeiből. A legtöbb jel, mint például a beszéd, vagy zaj legalább részben véletlenszerű. A véletlenszerűség elmélete az információ és a zaj elméletével teljesen összekapcsolódott. Valóban, a véletlenszerű jelek feldolgozásakor a munka nagy részét jelenti az információ kinyerése a zajosan vett jelből. Ha egy jel alkalmas információhordozásra, akkor valamilyen fokú véletlenszerűségnek kell benne lennie: a kiszámítható jelek nem hordoznak információt. Ezért a jel véletlenszerű

része lehet a jel információtartalma, vagy zaj, vagy a kettő keveréke, tehát információ és zaj. Bár a véletlenszerű jelek nem teljesen kiszámíthatók, mégis gyakran mutatnak egy sor olyan statisztikailag jellemző értéket, mint például a maximum, a minimum, a középérték, a felező, a szórásnégyzet, és a teljesítményspektrum. A véletlenszerű jel leírható a statisztikája szempontjából, és a legteljesebben a valószínűségi modell szempontjából, melyből az összes statisztikája számítható.

#### 2.2.1.1.1. Sztochasztikus folyamatok

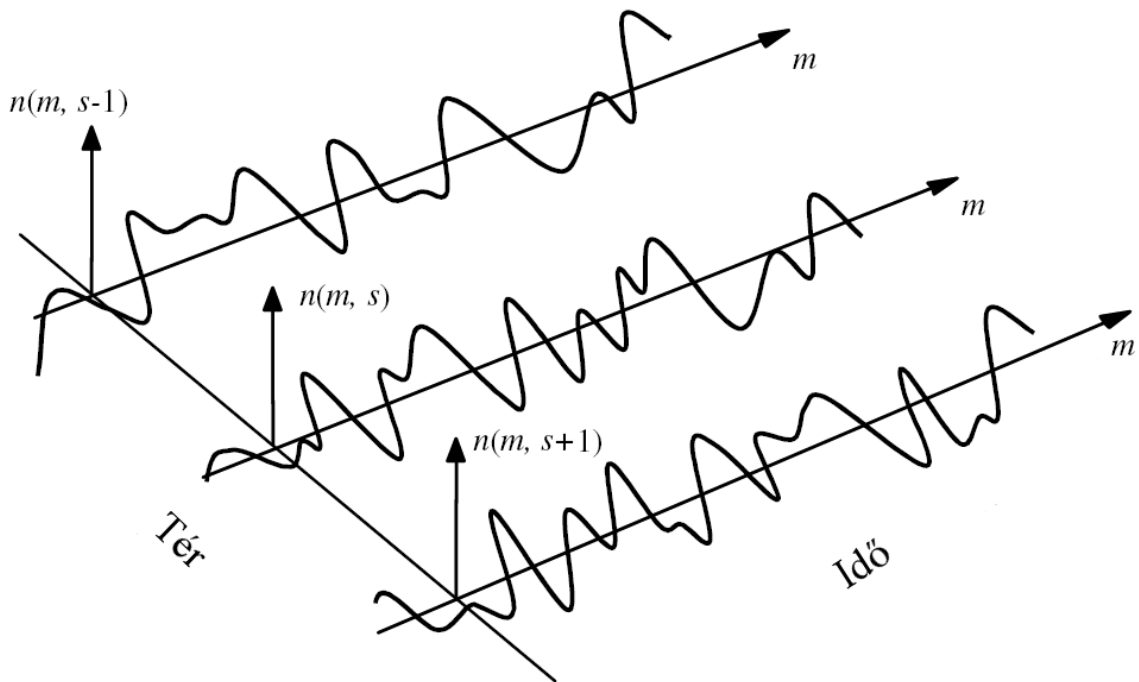
A „sztochasztikus eljárás” szakkifejezést általában a véletlenszerű folyamatok leírására használjuk, melyek folyamatos jeleket állítanak elő, mint például az emberi beszéd, vagy zaj. A jelfeldolgozási terminológiában egy sztochasztikus eljárás egy valószínűségi modell, a véletlenszerű jelek egy osztályára, például Gauss módszer, Markov módszer, Poisson módszer. Egy klasszikus példa a sztochasztikus eljárásra a részecskék úgynevezett Brown mozgása folyadékokban. A részecskék a folyadék belsejében véletlenszerűen mozognak a folyadékrészecskék „bombázása” következtében. Minden részecske véletlenszerű mozgása a sztochasztikus módszer egy egyszerű elképzelése. Minden részecske folyadékban történő mozgása segíti a megértést, vagy a módszer különféle megvalósításainak terét.

#### 2.2.1.1.2. Véletlen folyamatok együttese vagy tere

A véletlenszerű folyamatok összes elképzelésének gyűjteménye a módszerek együttese vagy tere néven ismert. Példaként tekintsünk egy véletlenszerű zajt egy telekommunikációs rendszerben ahogyan azt a 11. ábra is mutatja. A zaj minden telefonvonalon időben véletlenszerűen ingadozik, és jelölhető az  $n(m,s)$  jelöléssel, ahol  $m$  a diszkrét idő mérőszám, és  $s$  jelöli a sorindexet. A zajok gyűjteménye a különböző vonalakon a zaj együttesét (vagy terét) alkotják, melyet a  $N(m)=\{n(m,s)\}$  jelöléssel jelölünk, ahol  $n(m,s)$  az  $N(m)$  zaj realizációját jelöli az  $s$  vonalon.

#### 2.2.1.2. Valószínűségi modellek

A valószínűségi modellek a véletlenszerű jelek legteljesebb matematikai leírását adják. Egy  $m$  rögzített időpillanatban egy véletlenszerű jeltől gyűjtött minta  $\{x(m,s)\}$  egy valószínűségi változó, mely különböző értékeket vesz fel az  $s$  térben. Az alapvető különbség a véletlenszerű változó és a véletlenszerű folyamat között, hogy az utóbbi idősorozatot generál. Ezért a véletlenszerű változók modellezésére használt modellek használhatók a véletlenszerű folyamatokra is.



11. ábra A véletlen zaj három realizációja a zajok terében

A valószínűségi változó tere az összes érték gyűjteménye, vagy eredmény, amit a változó felvehet. A valószínűségi változó tere részekre osztható, néhány kritérium szerint, néhány altérre. Az altér jelértékek egy gyűjteménye, melyek egy közös tulajdonsággal rendelkeznek, mint egy csoportba összezsúfolt minta, vagy a minták gyűjteménye azok amplitúdójával az adott sávon belül. Minden alteret egy eseménynek hívunk, és az esemény bekövetkezésének valószínűsége  $A$ , és  $P(A)$  az a hányados, mely az  $A$  térből származó megfigyelések eredményeinek számát  $N_A$ -t

elosztjuk az összes megfigyelés számával:  $P(A) = \frac{N_A}{\sum_i N_i}$ . Az iménti egyenletből evi-

dens, hogy a kísérletben résztvevő összes lehetséges esemény valószínűségének összege egy.

#### 2.2.1.2.1. Valószínűségi változó eloszlása, eloszlásfüggvény

Legyen adott egy  $(\Omega, \mathbf{A}, P)$  valószínűségi mező, legyen értelmezve egy  $\xi: \Omega \rightarrow \mathbf{R}$  függvény, melyre teljesül, hogy  $\{\omega \mid \xi(\omega) = x\} \in \mathbf{A}$  minden  $x \in \mathbf{R}$  esetén. ( $\xi$  az elemi események halmazát képezi le  $\mathbf{R}$ -be, értékkészlete diszkrét értékekből áll). Ekkor  $\xi$ -t diszkrét valószínűségi változónak nevezzük.  $\xi$  értékkészlete  $\{x_1, x_2, \dots\}$  megszámlálható.  $\xi$  eloszlása:  $\{p_i = P(\xi = x_i) = P(\omega \mid \xi(\omega) = x_i) \mid i=1, 2, \dots\}$  valószínűségek halmaza, annyi elemű, ahány elemű az értékkészlet, elemei mutatják a diszkrét valószínűségi változó egyes értékeinek felvételi valószínűségét. Az értékkészlet elemei nemnegatív számok, melyek összege 1.



Két diszkrét valószínűségi változó ( $\xi$  és  $\eta$ ) *együttes eloszlásán* a következő eloszlást értjük (amennyiben  $\xi$  értékészlete  $\{x_1, x_2, \dots\}$  eloszlása  $\{P(\xi=x_k) (k=1, 2, \dots)\}$ ,  $\eta$  értékészlete  $\{y_1, y_2, \dots\}$  eloszlása  $\{P(\eta=y_j) (j=1, 2, \dots)\}$ ):  $\{P(\xi=x_k \text{ és } \eta=y_j) (k,j=1, 2, \dots)\}$ .

Az együttes eloszlás elemei nemnegatívak, az elemek összege 1. Az együttes eloszlásban szereplő események teljes eseményrendszer alkotnak, ezért az előző állítás igaz.

A peremeloszlás két diszkrét valószínűségi változó együttes eloszlásából meghatározható az alkotó diszkrét valószínűségi változók eloszlása. Valamelyik diszkrét valószínűségi változó eloszlása szerint összegezve az együttes eloszlást, a másik diszkrét valószínűségi változó eloszlásához jutunk.

Diszkrét valószínűségi *változók függetlenek*, ha minden  $j = 1, 2, \dots$  és minden  $k = 1, 2, \dots$  esetén  $P(\xi=x_k \text{ és } \eta=y_j)=P(\xi=x_k)P(\eta=y_j)$ , azaz együttes eloszlásuk minden eleme szétbomlik az eloszlások megfelelő elemeinek szorzatára. Két diszkrét valószínűségi változó eloszlása pontosan akkor határozza meg együttes eloszlásukat, ha függetlenek. A konstans értékű valószínűségi változó minden más diszkrét valószínűségi változótól független.

Adott valószínűségi mezőben lévő  $B_i (i=1, 2, \dots)$  teljes eseményrendszer esetén ha ezek között nincs 0 valószínűségű esemény, akkor tetszőleges nem 0 valószínűségű  $A$  eseményre  $P(A | B) = \frac{P(A | B_i) \cdot P(B_i)}{\sum_i (P(A | B_i) \cdot P(B_i))}$ . Ez Bayes tétele.

Egy  $\xi$  valószínűségi változó *eloszlásfüggvénye*:  $F: \mathbf{R} \rightarrow [0, 1]$ ,  $F(x)=P(\xi < x)$ . Valószínűségi változó eloszlásfüggvénye balról folytonos, monoton nemcsökkenő,  $-\infty$ -ben 0-hoz tart,  $+\infty$ -ben 1-hez tart. Bármely, a valószínűségi változó eloszlásfüggvényeire vonatkozó fenti feltételeket kielégítő függvényhez létezik olyan valószínűségi változó, melynek eloszlásfüggvénye az adott függvény. A diszkrét valószínűségi változónak is létezik eloszlásfüggvénye, ez „lépcsős” alakú.  $\xi$  valószínűségi változó esetén  $P(a \leq \xi < b) = P(\xi < b) - P(\xi < a) = F(b) - F(a)$ , ha  $a < b$ .  $\xi$  és  $\eta$  *valószínűségi változók együttes eloszlásfüggvénye*  $F(x,y)=P(\xi < x \text{ és } \eta < y) (x,y \in \mathbf{R})$ .  $\xi$  és  $\eta$  *valószínűségi változók függetlenek*, ha együttes eloszlásfüggvényük megegyezik eloszlásfüggvényeik szorzatával. Valószínűségi változók együttes eloszlásfüggvénye mindkét változóiban monoton nemcsökkenő, balról folytonos,  $\lim_{x \rightarrow -\infty} F(x,y) = \lim_{y \rightarrow -\infty} F(x,y) = 0$ ,  $\lim_{y \rightarrow \infty, x \rightarrow \infty} F(x,y) = 1$ ; ha  $a_1 < b_1$  és  $a_2 < b_2$ , akkor  $F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) - F(a_1, a_2) \geq 0$ . Bármely, a valószínűségi változók együttes eloszlásfüggvényére vonatkozó ezen feltételeket kielégítő függvényhez létezik két olyan valószínűségi változó, melyek-

nek együttes eloszlásfüggvénye az adott függvény. Az utolsó feltétel  $P(\xi \in [a_1, b_1] \text{ és } \eta \in [a_2, b_2]) \geq 0$ -t jelenti.

### 2.2.1.2.2. Sűrűségfüggvény

Egy adott eseménytéren értelmezett  $\xi: \Omega \rightarrow \mathbf{R}$  függvény *valószínűségi változó*, ha  $\{\omega \mid \xi(\omega) < x\} \in \mathbf{A}$  bármely  $x \in \mathbf{R}$  esetén. A diszkrét valószínűségi változó teljesíti az iménti definíció követelményeit. A definíció a diszkrét valószínűségi változók általánosítása: folytonos esetben  $P(\xi=x)=0$  minden  $x \in \mathbf{R}$  esetén. Valószínűségi változót (ha nem hangsúlyozottan diszkrét) a definíció szerinti értelemben kell érteni.

Egy  $\xi$  valószínűségi változó *folytonos*, ha eloszlásfüggvénye abszolút folytonos. Folytonos valószínűségi változó esetén  $P(\xi=x)=0$  bármely  $x \in \mathbf{R}$  esetén.

Belátható, hogy  $0 \leq P(\xi=x) = \lim_{\Delta x \rightarrow 0} P(x \leq \xi < x + \Delta x) = \lim_{\Delta x \rightarrow 0} (F(x + \Delta x) - F(x)) \leq \lim_{\Delta x \rightarrow 0} ((F(x) + K\Delta x) - F(x)) = \lim_{\Delta x \rightarrow 0} (K\Delta x) = 0$ .

Bármely  $\xi$  valószínűségi változó esetén bármely  $x \in \mathbf{R}$ -re  $P(a \leq \xi \leq b) = F(b+0) - F(a)$ ,  $P(a < \xi < b) = F(b) - F(a-0)$ ,  $P(a < \xi \leq b) = F(b+0) - F(a-0)$  teljesül. Folytonos esetben ezek mind egyenlők  $P(a \leq \xi < b)$ -vel. Folytonos  $\xi$  valószínűségi változó *sűrűségfüggvénye*

$f: \mathbf{R} \rightarrow \mathbf{R}$ , ha  $\int_{-\infty}^x f(t) dt = F(x)$  teljesül minden  $x \in \mathbf{R}$  esetén. Folytonos valószínűségi

változó sűrűségfüggvényére és eloszlásfüggvényére teljesül az  $F'(x) = f(x)$  összefüggés. Diszkrét valószínűségi változónak nincs sűrűségfüggvénye.

A sűrűségfüggvény nemnegatív, az egész számegegyenesen vett integrálja 1. Bármely, a valószínűségi változó sűrűségfüggvényére vonatkozó, ezen feltételeket kielégítő függvényhez létezik olyan valószínűségi változó, melynek sűrűségfüggvénye az adott függvény. Diszkrét esetben hasonló feltétel a  $p_i \geq 0$ ,  $\sum p_i = 1$  tulajdonságok: az eloszlás játssza a sűrűségfüggvény szerepét is.

$\xi$  és  $\eta$  abszolút folytonos *valószínűségi változók együttes sűrűségfüggvénye*

$f(u, v)$ , ha rá teljesül az  $\int_{-\infty}^x \int_{-\infty}^y f(u, v) dv du = F_{\xi, \eta}(x, y)$  egyenlőség.

Az együttes sűrűségfüggvény megegyezik az együttes eloszlásfüggvény második parciális deriváltjával (először az első, majd a második változó szerint deriválva). Abszolút folytonos  $\xi$  és  $\eta$  valószínűségi változók esetén függetlenségük egyenértékű  $f_{\xi, \eta}(u, v) = f_{\xi}(u)f_{\eta}(v)$  teljesülésével. Az együttes sűrűségfüggvény nemnegatív, az egész valós számsíkon vett integrálja 1. Az együttes sűrűségfüggvényből meghatározható az alkotó valószínűségi változók sűrűségfüggvénye. Az együttes sűrűség-

függvényt az egyik változó szerint az egész valós számegyenesen integrálva a másik valószínűségi változó sűrűségfüggvényét kapjuk.

### 2.2.1.3. Állandó és nem állandó véletlen folyamatok

Bár a jel  $x(m)$ -mel jelölt amplitúdója  $m$  periódussal hullámzik, a jel által generált folyamat jellemzője lehet időben állandó vagy időben változó. Nem állandó folyamat például a beszéd, melynek hangossága és spektrális jellemzői, összetevői folyamatosan változnak, ahogyan a beszélő különféle hangokat képez. A folyamat állandónak tekinthető, ha a valószínűségi modellbeli jellemzői idő invariánsak, ellenkező esetben nem állandó – mindkettőt a 12. ábra szemlélteti. Az állandó tulajdonság magába foglalja, hogy a jel összes paramétere, úgy mint középérték, variancia, energiaspektrum összetétele, és a folyamat magasabb rendű momentumai idő invariánsak. A gyakorlatban az állandóságnak különféle fokai ismeretesek. Van olyan, hogy a statisztikák egy halmaza állandó, a többi pedig nem. Például lehet egy véletlen folyamatnak idő invariáns középértéke, de az idő függvényében változó energiája.



12. ábra Kvázistacionárius és nemstacionárius beszédsgemensek

### 2.2.1.3.1. Szigorú értelemben vett állandó folyamatok

Egy  $X(m)$  véletlenszerű folyamat szigorú értelemben véve stacionárius, ha minden megoszlása, és statisztikai paramétere idő invariáns. A szigorú értelemben vett állandóság magában foglalja, hogy az  $n$ -edik rendű eloszlása transláció invariáns minden  $n = 1, 2, \dots$ , vagyis:

$$P[x(m_1) \leq x_1, x(m_2) \leq x_2, \dots, x(m_n) \leq x_n] = P[x(m_1 + \tau) \leq x_1, x(m_2 + \tau) \leq x_2, \dots, x(m_n + \tau) \leq x_n]$$

Az egyenletből a szigorú értelemben vett állandó folyamat statisztikája beleértve a középértéket, a korrelációt és az energiaspektrumot, idő invariáns lesz, ezért:  $E[x(m)] = \mu_x$ ,  $E[x(m)x(m+k)] = r_{xx}(k)$  és  $E[|X(f, m)|^2] = E[|X(f)|^2] = P_{xx}(f)$ , ahol  $\mu_{xx}$ ,  $r_{xx}(m)$  és  $P_{xx}(f)$  rendre az  $x(m)$  jel középértéke, az autokorrelációja és az energiaspektruma, és  $X(f, m)$  jelöli  $x(m)$  frekvencia-idő spektrumát.

### 2.2.1.3.2. Nem szigorú értelemben vett állandó folyamatok

Míg a szigorú értelemben vett állandó (stacionárius) folyamatok esetében a minden statisztikának idő-invariánsnak kell lennie, a nem szigorú értelemben vett stacionárius folyamatokra nem vonatkozik az összes korlátozás. Egy ilyen folyamat középértéke és az autokorrelációs függvénye idő invariáns:  $E[x(m)] = \mu_x$ ,  $E[x(m)x(m+k)] = r_{xx}(k)$ . A két folyamat definíciójából kiderül, hogy minden szigorú értelemben vett stacionárius folyamat nem szigorú értelemben is az, viszont ennek a fordítottja nem igaz.

### 2.2.1.3.3. Nem stacionárius folyamatok

Egy véletlenszerű folyamat nem stacionárius, ha eloszlása vagy statisztikája időben változik. A legtöbb sztochasztikus folyamat, mint például a videojelek, audiojelek, gazdasági adatok, meteorológiai adatok, stb., nem stacionárius, mert egy-egy olyan rendszer generálja őket, melynek környezeti paraméterei időben változnak. Például a beszéd nem stacionárius, mert egy időben változó artikulációs rendszer állítja elő. A beszéd hangerőssége, frekvencia-összetétele időben változó, mely változás néha teljesen váratlan. Az időben változó folyamatok modellezhetők stacionárius véletlenszerű folyamatok egy kombinációjával.

### 2.2.1.4. Véletlen folyamatok várható értéke

A várható érték központi szerepet játszik a jelek modellezésében és feldolgozásában. Továbbá a véletlenszerű folyamatok valószínűségi modelljei általában a várható érték függvényeként kerülnek kifejezésre. Például a normális eloszlás el-

oszlásfüggvénye a folyamat középértékét és kovarianciáját tekintve egy exponenciális függvényként definiálható, a Poisson eloszlás eloszlásfüggvénye pedig a folyamat középértékével definiálható. A jelfeldolgozási alkalmazásokban mindig megvan a jel megfelelő statisztikai modellje – például a normális eloszlás eloszlásfüggvénye –, és a modell kiegészítésére szükség van a várható paraméterek értékére. Azonkívül sok jelfeldolgozási algoritmusban, mint például a spektrális kivonásban a zajcsökkentésre, vagy a lineáris előrejelzésben, lényegében amire szükségünk van az a jel középértékének vagy a korrelációs függvényének kiszámítása. Egy véletlenszerű folyamat  $h(X(m_1), X(m_2), \dots, X(m_M))$  függvényének várható értékének kiszámítása a következőképpen definiálható:

$$E[h(X(m_1), \dots, X(m_M))] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} h(x_1, \dots, x_M) f_{X(m_1) \dots X(m_M)}(x_1, \dots, x_M) dx_1 \dots dx_M$$

A legfontosabb és széles körben használt várható értékek a középérték, a korreláció, a kovariancia és a energiaspektrum.

#### 2.2.1.4.1. A középérték

A egy jelsorozat középértéke nagyon fontos szerepet játszik a jelfeldolgozásban és paraméterkiszámításban a zajos jelek esetében. Például egy zajos jel esetében az optimális lineáris kiszámítási módszer az interpoláció a középérték és a megfigyelt jel zajjal terhelt értéke között. Egy  $[X(m_1), \dots, X(m_M)]$  véletlenszerű vektor középértéke alatt értjük a teljes jelfolyam alatt mutatott átlagértéket, mely a következőképpen definiálható:

$$E[X(m_1), \dots, X(m_M)] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (x_1, \dots, x_M) f_{X(m_1) \dots X(m_M)}(x_1, \dots, x_M) dx_1 \dots dx_M$$

#### 2.2.1.4.2. Autokorreláció

A korrelációs függvény és annak Fourier transzformációja, az energiaspektrum sűrűsége minták és struktúrák modellezésére és felismerésére használatos a jelfeldolgozások során. A kölcsönösségi együtthatók központi szerepet játszanak a jelfeldolgozásban, a telekommunikációs rendszerekben, beleértve a prediktív kódolókat, kiegyenlítőket, digitális dekódereket, késleltetés kiszámítókat, osztályozókat és jelvisszaállító rendszereket. Egy  $X(m)$  véletlenszerű folyamat autokorrelációs függvénye, melyet  $r_{xx}(m_1, m_2)$ -vel jelölünk, a következőképpen definiálható:

$$r_{xx}(m_1, m_2) = E[x(m_1)x(m_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(m_1)x(m_2) f_{X(m_1), X(m_2)}(x(m_1), x(m_2)) dx(m_1) dx(m_2)$$

Az  $r_{xx}(m_1, m_2)$ -vel jelölt autokorrelációs függvény az  $X$  folyamat végeredményének  $m_1$  és  $m_2$  időpillanatbeli hasonlóságának mértékét vagy a kölcsönösségi relációját mutatja meg. Ha a véletlenszerű folyamat végeredménye az  $m_1$  és az  $m_2$  időpillanatbeli értéke nincs hordoz relációt, akkor  $X(m_1)$  és  $X(m_2)$  függetlennek vagy korrelálatlannak mondható, tehát  $r_{xx}(m_1, m_2) = 0$ . A nem szigorú értelemben vett stacionárius folyamatok esetében az autokorrelációs függvény idő invariáns és a két időpillanat között eltelt idő hosszának függvénye, vagyis  $m = m_1 - m_2$ :  $r_{xx}(m_1 + \tau, m_2 + \tau) = r_{xx}(m_1, m_2) = r_{xx}(m_1 - m_2) = r_{xx}(m)$ . Egy valós értékű nem szigorú értelemben vett stacionárius folyamat autokorrelációs függvénye szimmetrikus a következő tulajdonságokkal:  $r_{xx}(-m) = r_{xx}(m)$  és  $r_{xx}(m) \leq r_{xx}(0)$ . Ez esetben az  $r_{xx}(0)$  jelenti a nulla középértékű jel energiáját.

#### 2.2.1.4.3. Autokovariancia

Az autokovariancia függvény egy  $X(m)$  véletlenszerű folyamat  $c_{xx}(m_1, m_2)$  jelöléssel ellátott szórásának vagy szóródásának mérésére szolgál a folyamat középértéke körül és a következőképpen definiálható:

$$c_{xx}(m_1, m_2) = E[(x(m_1) - \mu_x(m_1))(x(m_2) - \mu_x(m_2))] = r_{xx}(m_1, m_2) - \mu_x(m_1)\mu_x(m_2),$$

ahol a  $\mu_x(m)$  jelenti az  $X(m)$  véletlenszerű folyamat középértékét. Nulla középértékű folyamatok esetében az autokovariancia és az autokorreláció függvények azonosak. A  $c_{xx}(m_1, m_2)$  a folyamat varianciáját (szórásnégyzetét) jelenti. Stacionárius folyamatok esetében az autokovariancia függvény az előző egyenletből a következőképpen alakul:  $c_{xx}(m_1, m_2) = c_{xx}(m_1 - m_2) = r_{xx}(m_1 - m_2) - \mu_x^2$ .

#### 2.2.1.4.4. Spektrális energiasűrűség

Egy véletlenszerű folyamat energiájának spektrális sűrűségfüggvénye, más néven energiaspektruma megadja a jel egyes frekvenciákhoz tartozó energiáinak eloszlását a teljes spektrumon. A nem szigorú értelemben vett stacionárius folyamat  $X(m)$  energiaspektruma a Wiener–Khinchin tétel alapján úgy definiálható, hogy az autokorrelációs függvény Fourier transzformáltja és a következőképpen írható fel:  $P_{XX}(f) = E[X(f)X^*(f)] = \sum_{m=-\infty}^{\infty} r_{xx}(m)e^{-j2\pi fm}$ , ahol  $r_{xx}(m)$  és  $P_{XX}(f)$  rendre az  $x(m)$  autokorrelációs függvénye és energiaspektruma, és  $f$  pedig a frekvenciaváltozó. Egy valós értékű stacionárius folyamat autokorrelációs függvénye szimmetrikus, és az energiaspektruma a következőképpen írható fel:

$$P_{XX}(f) = r_{xx}(0) + \sum_{m=1}^{\infty} 2r_{xx}(m) \cos(2\pi fm)$$

A spektrális sűrűségfüggvény egy nemnegatív valós értékészletű felvevő függvény, megadása pedig watt per herzben  $\left[\frac{\text{W}}{\text{Hz}}\right]$  történik. Az autokorrelációs so-

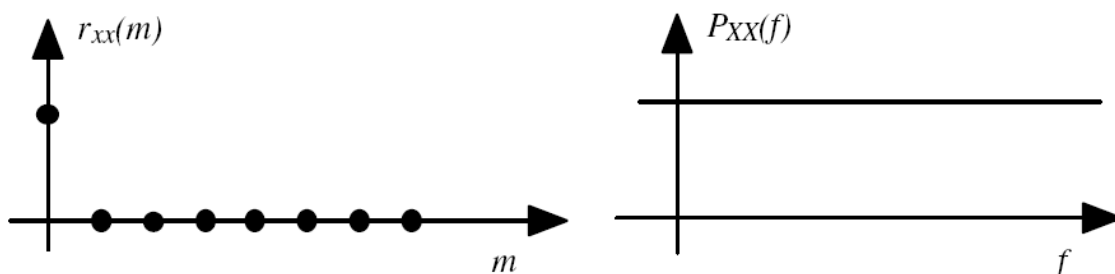
rozat a véletlenszerű folyamat energiaspektrumából az iménti egyenlet felhasználásával kapható meg az inverz Fourier transzformáció alkalmazásával a követke-

zőképpen:  $r_{xx}(m) = \int_{-\frac{1}{2}}^{\frac{1}{2}} P_{XX}(f) e^{j2\pi fm} df$ . Az autokorreláció és az energiaspektrum a

folyamat másodrendű statisztikai, rendre időben és frekvenciatartományban.

#### Példa – fehér zaj

A következő ábra a fehér zaj energiaspektrumát és autokorrelációját mutatja.



13. ábra A fehér zaj autokorrelációja és energiaspektruma

Egy olyan zajt, amely korrelálatlan, egymástól független mintákból áll, fehér zajnak nevezzünk. Egy stacionárius fehér zaj  $n(m)$  autokorrelációja definiálható:

$r_{nn}(k) = E[n(m)n(m+k)] = \begin{cases} \text{Zajenergia} & k = 0 \\ 0 & k \neq 0 \end{cases}$ . Ez az egyenlet egy korrelálatlan fehér

zaj definíciója szerinti matematikai állítás. A frekvenciatartományra ekvivalens leírás adható meg, az  $r_{nn}(k)$  Fourier transzformációjával a következőképpen:

$P_{NN}(f) = \sum_{k=-\infty}^{\infty} r_{nn}(k) e^{-j2\pi fk} = r_{nn}(0) = \text{zajenergia}$ . A stacionárius fehér zaj energiaspekt-

ruma azonosan terjed szét a sáv teljes szélességében és időben. A fehér zaj az egyik legnehezebben eltávolítható zajforma, mert nem határozható meg a szerkezete sem időben, sem a frekvenciasávok szerint.

#### Példa – impulzív zaj

Az impulzív zaj egy véletlenszerű kétállapotú („be/ki”) impulzussorozatból áll, véletlenszerű amplitúdóval és időben véletlenszerű előfordulással. Egy véletlenszerű impulzív zaj sorozat  $n_i(m)$  egy kétállapotú véletlenszerű amplitúdó modulált sorozatként modellezhető, a következő szerint:  $n_i(m) = n(m)b(m)$ , ahol  $b(m)$  egy

kétállapotú sorozat, mely egy impulzus jelenlétét vagy hiányát jelzi, és  $n(m)$  pedig egy véletlenszerű zaj. Tegyük fel, hogy az impulzív zaj egy korrelálatlan folyamat, így az impulzív zaj, mint egy kétállapotú folyamat autokorrelációja a következőképpen definiálható:  $r_m(k, m) = E[n_i(m)n_i(m+k)] = \sigma_n^2 \delta(k)b(m)$ , ahol  $\sigma_n^2$  a zaj varianciája. Ebben az egyenletben az autokorrelációs függvény egy kétállapotú függvényként lett kifejezve, mely az impulzív zaj  $m$  időpillanatbeli be/ki állapotától függ. Az impulzív zaj sorozat energiaspektruma az imént leírt autokorrelációs függvény Fourier transzformáltjaként kapható meg a következő összefüggéssel:  $P_{NN}(f, m) = \sigma_n^2 b(m)$ .

#### 2.2.1.4.5. Két véletlenszerű folyamat együttes statisztikája

A jelfeldolgozás során gyakran találkozunk azzal a problémával, például egy több érzékelőből álló rendszer esetében, hogy nem csak egy véletlenszerű folyamat van jelen. Az együttes statisztika és együttes eloszlás leírja a két vagy több folyamat kapcsolatát egymással. Két diszkrét időbeosztással mintavételezett véletlenszerű folyamat  $x(m)$  és  $y(n)$  együttes eloszlásfüggvénye a következő jelöléssel adható meg:  $f_{X(m_1)\dots X(m_M), Y(n_1)\dots Y(n_N)}(x_1, \dots, x_M, y_1, \dots, y_N)$ . Ha a két véletlenszerű folyamat  $X(m)$  és  $Y(n)$  korrelálatlanok, az együttes eloszlásfüggvényük a két eloszlásfüggvény szorzataként adható meg a következő felírással:

$$f_{X(m_1)\dots X(m_M), Y(n_1)\dots Y(n_N)}(x_1, \dots, x_M, y_1, \dots, y_N) = f_{X(m_1)\dots X(m_M)}(x_1, \dots, x_M) f_{Y(n_1)\dots Y(n_N)}(y_1, \dots, y_N)$$

#### 2.2.1.4.6. Kereszt-korreláció és kereszt-kovariancia

Két véletlenszerű folyamat kereszt-korrelációja az alábbiak szerint adható

$$\text{meg: } r_{xy}(m_1, m_2) = E[(x(m_1)y(m_2))] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(m_1)y(m_2)f_{X(m_1)Y(m_2)}(x(m_1), y(m_2))dx(m_1)dy(m_2)$$

Egy  $r_{xy}(m_1, m_2)$  nem szigorú értelemben vett stacionárius folyamat esetében a korrelációs függvény csak az  $m_1$  és  $m_2$  időpillanatok különbségétől függ, vagyis  $m = m_1 - m_2$ , tehát:  $r_{xy}(m_1 + \tau, m_2 + \tau) = r_{xy}(m_1, m_2) = r_{xy}(m_1 - m_2) = r_{xy}(m)$ .

A kereszt-kovariancia függvény a következőképpen definiálható:  $c_{xy}(m_1, m_2) = E[(x(m_1) - \mu_x(m_1))(y(m_2) - \mu_y(m_2))] = r_{xy}(m_1, m_2) - \mu_x(m_1)\mu_y(m_2)$ . Nulla középértékkel rendelkező folyamatok esetében a kereszt-korreláció és a kereszt-kovariancia függvények azonosak. Nem szigorú értelemben vett stacionárius folyamatok esetében a kereszt-kovariancia függvény az előző egyenlet alapján:  $c_{xy}(m_1, m_2) = c_{xy}(m_1 - m_2) = r_{xy}(m_1 - m_2) - \mu_x\mu_y$ .



## 2.2.1.4.7. Kereszt-energiáspektrum eloszlás és koherencia

Két véletlenszerű folyamat  $X(m)$  és  $Y(m)$  kereszt-energiáspektrum eloszlása a két folyamat kereszt-korrelációs függvényének Fourier transzformációjával definiálható:

$$P_{XY}(f) = E[X(f)Y^*(f)] = \sum_{m=-\infty}^{\infty} r_{xy}(m)e^{-j2\pi fm}.$$

Ahogy a kereszt-korreláció a kereszt-energiáspektrum eloszlása is a két jel hasonlóságának mértékét mutatja, vagy a két folyamat spektrális koherenciáját, mely a kereszt-energiáspektrum eloszlásának normalizált formulája és a következőképpen definiált:

$$C_{XY}(f) = \frac{P_{XY}(f)}{\sqrt{P_{XX}(f)P_{YY}(f)}}.$$

A koherencia függvény olyan alkalmazásokban használatos, mint az időkésettetés kiszámítása és a jel-zaj arány meghatározása (mérése).

## 2.2.2. A rejtett Markov modell

A rejtett Markov modelleket a véletlenszerű nem stacionárius jelek statisztikai modellezésére használhatjuk hatékonyan. Az ilyen jelek csoportjába tartozik az emberi beszéd vagy az időben változó zaj. A HMM egy Markov láncot ad meg, mely az időbeli változást vagy a terek változását állapotfüggő stacionárius alfolyamatokra bontja. Egy HMM lényegében egy véges állapothalmazzal rendelkező Bayes folyamat, Markov előzménnyel, az egyes állapotok között átmenetek modellezésére, és egy halmaz valószínűségi eloszlásfüggvény a jel véletlenszerű variációinak modellezésére minden állapoton belül.

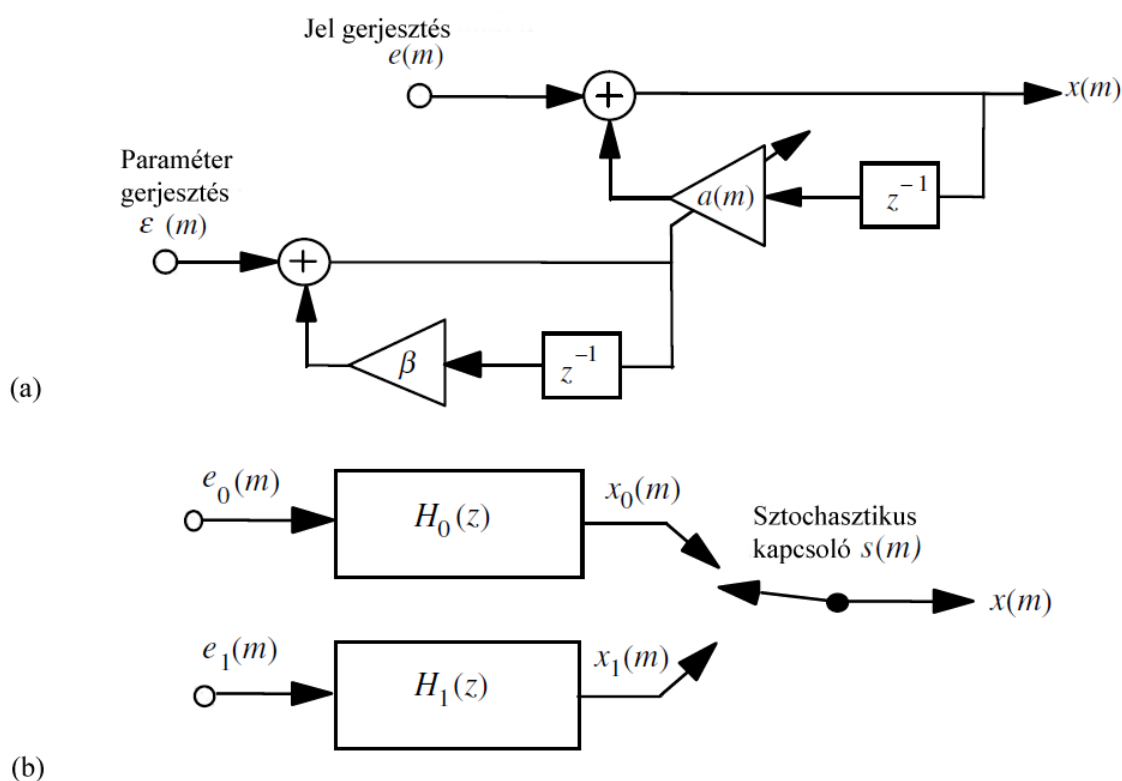
2.2.2.1. Nem stacionárius folyamatok statisztikai modellje

Egy nem stacionárius folyamat definiálható oly módon, hogy egy olyan folyamat, melynek statisztikai paraméterei az idővel változnak. A legtöbb „természetesen előállított” jel, mint például a hangjelek, képek, biomedikai és szeizmikus jelek nem stacionárius jelek, mivel az azokat előállító rendszerek paraméterei és a környezet, melyben terjednek változnak az idő múlásával.

A nem stacionárius jelek egy kétrétegű sztochasztikus folyamatként modellezhetők, egy rejtett folyamattal az észlelhető jel statisztikai paramétereinek időbeli változásainak kontrollálására. Általában a nem stacionárius folyamatok két alapvető nagy csoportba sorolhatók:

- a.) folytonosan változó állapotú folyamatok
- b.) véges sok állapotú folyamatok

A folytonosan változó állapotú folyamatok definíciós szerint az időben folytonosan változó jelek statisztikájának alapjául szolgálnak. Ilyenek például az audiojelek, a beszéd és a zene, melyeknek az energiája és spektrális összetétele az idővel változik. Véges sok állapottal rendelkező folyamat az olyan statisztikai jellemzőkkel rendelkező folyamat, mely képes váltani a véges sok stacionárius vagy nem stacionárius állapot között. Például az impulzív zaj egy ilyen kétállapotú folyamat. Folytonosan változó folyamatok közelíthetők hozzájuk illő véges sok állapotú folyamattal. A következő ábra egy nem stacionárius autoregresszív (AR) folyamatot mutat be. Ez a folyamat a jel paramétereire egy rejtett stacionárius AR modell és magára a jelre egy időben változó AR modell kombinációját alkalmazza. A



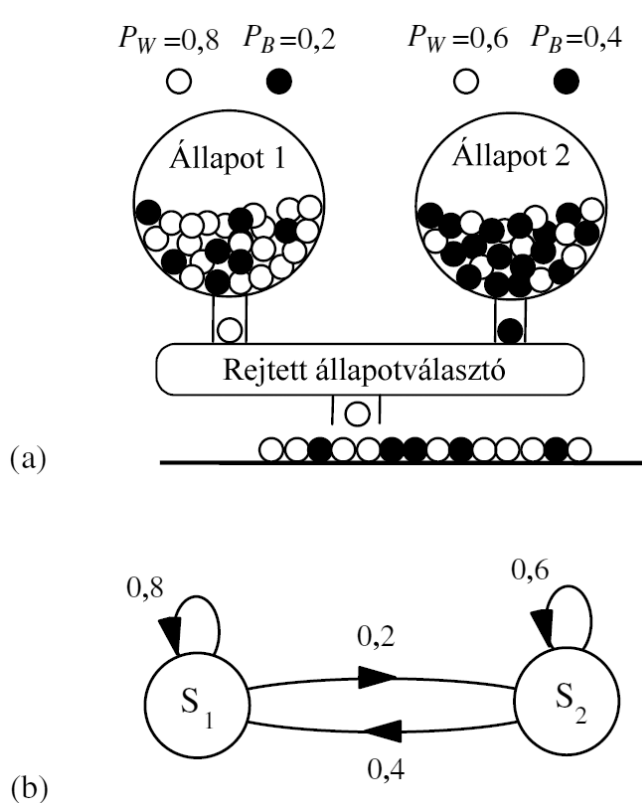
14. ábra Folytonosan változó (a) és kétállapotú (b) AR folyamat

rejtett modell vezérli a nem stacionárius AR modell időben változó paramétereit. Ehhez a modellhez a megfigyelt jel egyenlete és a paraméter állapotegyenlete a következő módon írható fel: a megfigyelt jel  $x(m) = a(m)x(m-1) + e(m)$  és a rejtett állapot  $a(m) = \beta a(m-1) + \varepsilon(m)$ , ahol  $a(m)$  a megfigyelt AR folyamat időben változó együtthatója és  $\beta$  pedig a rejtett állapotirányító folyamat együtthatója. A 14. (b) ábra egy egyszerű példát mutat a véges sok állapotú nem stacionárius folyamat kétállapotú autoregresszív megvalósítására, amelyben egy közvetlen kapcsoló választja ki minden időpillanatban a két AR modell közül az egyiket, melyet a

kimenetre csatol. Ehhez a modellhez a kimeneti jel  $x(m)$  a következő egyenlettel fejthető ki:  $x(m) = \bar{s}(m)x_0(m) + s(m)x_1(m)$ , ahol a kétállású kapcsoló  $s(m)$  kiválasztja a folyamat állapotát  $m$  időpillanatban, és  $\bar{s}(m)$  jelöli  $s(m)$  ellentettjét.

### 2.2.2.2. Rejtett Markov modellek

A rejtett Markov modell (HMM) egy kétrétegű véges sok állapotú folyamat, egy rejtett Markov folyamattal, mely irányítja az állapotok kiválasztását a figyelembe veendő folyamatban. Mint egy egyszerű illusztrációt figyeljük meg a következő ábrán látható kétállapotú Markov folyamatot, mely két tárolót mutat különböző keverékű fehér és fekete színű golyókkal. A fekete és fehér golyók előfordulásának valószínűségét jelöljük rendre  $P_B$  és  $P_W$  jelölésekkel, ahogyan az ábra mutatja. Tegyük fel, hogy folytonos idő intervallumban egy rejtett kiválasztó folyamat kivá-



15. ábra Kétrétegű véletlenszerű folyamat (a), és rejtett Markov modellje (b)

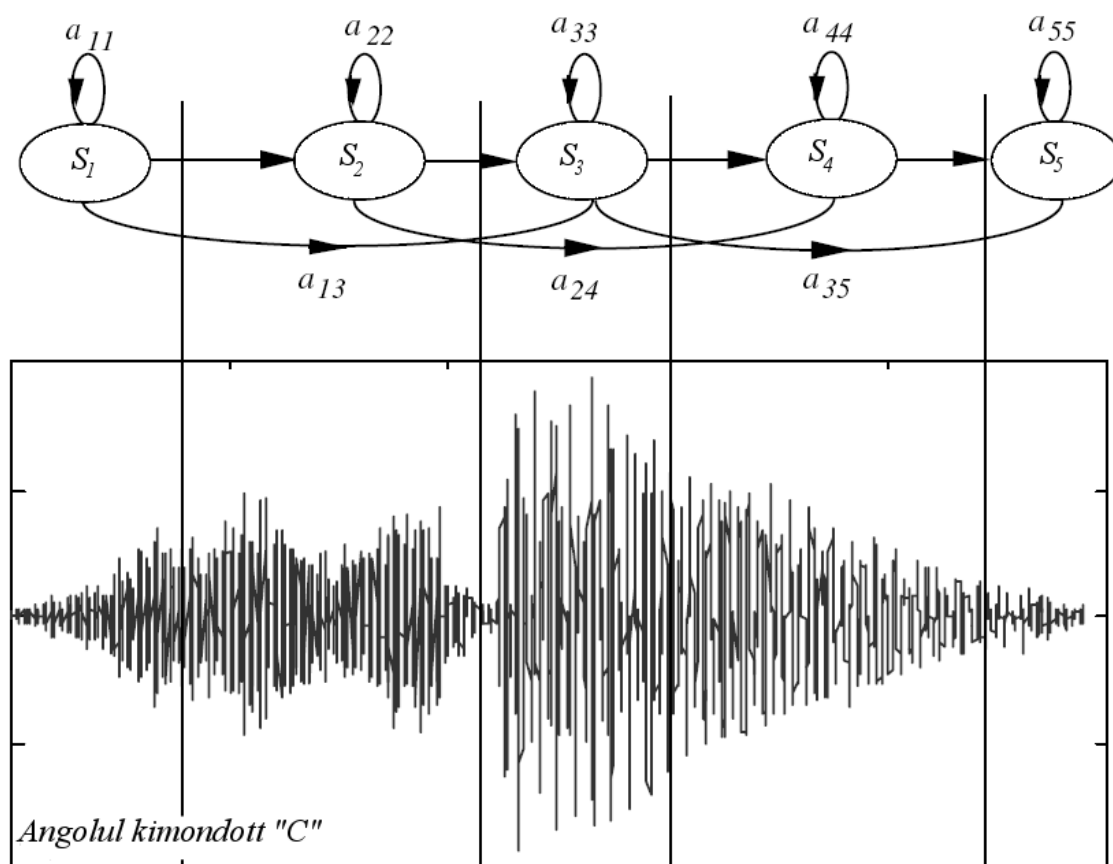
lasztja az egyik tartályt, hogy kiejtsen egy golyót. A golyót kiejtése után pótoljuk, tehát a fehér és fekete golyók számára nincs hatása a műveletnek. Minden tárolót úgy tekinthetünk, hogy az egy alapvető állapota a kimeneti folyamatnak. Most tegyük fel, hogy a rejtett tároló-kiválasztót a következő szabály irányítja: bármely időpillanatban, ha a kiválasztott tárolóból fehér golyót kapunk, akkor nem választunk másik tárolót, ha feketét akkor a másik tárolót választjuk. Ez egy példa a Markov folyamatra, mert a folyamat következő állapota függ a jelenlegi állapottól, ahogyan azt az ábra kétállapotú modellje is mutatja. Ebben a modellben a megfigyelhető végeredmény nem jelzi egyértelműen az alapul vett állapotot, mivel mindkét állapot (tároló) alkalmas minkét színű (fekete és fehér) golyó kibocsátására.

Általában a Markov modellnek  $N$  állapota van, és mindegyik állapot más szegmensét modellezi a vizsgált jelnek. A rejtett Markov modell alkalmazható egy

időben változó véletlenszerű folyamat modellezésére, mint egy  $N$  elemű stacionárius vagy kvázi-stacionárius elemi alfolyamat valószínűségi Markov lánc.

### 2.2.2.2.1. A rejtett Markov modellek fizikai értelmezése

A modell fizikai értelmezéséhez nézzük meg a következő, 16. ábrát. Az ábra balról jobbra mutatja egy angolul kiejtett „C”<sup>1</sup> hang modelljét és vele együtt a hang hullámformájának térképét is. Általában kétféle eltérés van a beszéd és az egyéb sztochasztikus jelek között: a spektrális összeállítás változása, és az időbeosztás vagy artikuláció változása. Egy rejtett Markov modellben ezek a változások állapot megjegyzéssel és az egyes állapotok közötti átmenetek valószínűségével modellezhetők. Használható módszer a rejtett Markov modell interpretálására és alkalmazására, hogy a modell minden állapotát úgy tekintjük, hogy a sztochasztikus folyamat egy szegmensét modellezi. Az ábrán az  $S_1$  állapot modellezi az első szeg-



16. ábra Ötállapotú balról–jobbra HMM beszédmodell

<sup>1</sup> Az angol abc 3. betűjének, a c-nek kiejtése a magyartól jelentősen eltér, fonetikusan a következőképpen írható le: [si:]

menet, a  $S_2$  a másodikat és így tovább. Minden állapotnak kell lennie egy olyan mechanizmusának, mely összeilleszti a szegmensek véletlenszerű variációinak különböző realizációit, melyet modelleznek. Az állapotátmenetek valószínűségei egy mechanizmust biztosítanak a különböző állapotok kapcsolatára, és a variációk modellezésére az időskála jelek alatti időtartamára minden állapotban. Például ha a beszéd egy szegmensének vége kinyúlt, köszönhetően a lassú artikulációnak, hozzáigazítható lesz több önmagába visszatérő átmenettel az állapotához, mely a szegmenst modellezi. Ellenben ha a szegmens elhallgatott a gyors beszédnek köszönhetően, akkor a következő állapotra ugrással illeszt ebben az esetben. Az állapot tapasztalati eloszlásfüggvényei modellezi a jel állapotokhoz rendelt szegmenseinek spektrális összeállításának valószínűségi eloszlását.

#### 2.2.2.2.2. A rejtett Markov modell paraméterei

A rejtett Markov modellnek a következő paraméterei vannak:

- Az állapotok száma  $N$ . Ezt általában beállítjuk az egymástól különböző vagy elemi sztochasztikus események számára egy jelfolyamatban. Például egy impulzív zajhoz hasonló kétállapotú folyamat modellezése esetén  $N$  értéke 2 lesz, míg szavankénti beszéd esetében  $N$  értékét 5 és 10 közé szokás választani.
- Állapot-átmeneti valószínűségi mátrix  $\mathbf{A} = \{a_{ij}, i, j = 1, \dots, N\}$ . Ez biztosít kapcsolatot vagy Markov hálózatot hoz létre az egyes állapotok között, és modellezi a variációkat az állapotokhoz rendelt jelek időtartama alatt. A balról–jobbra rejtett Markov modell esetében  $a_{ij} = 0$ , minden  $i > j$  esetén, ezért ez az  $\mathbf{A}$  mátrix egy felső háromszög mátrix.
- Állapot tapasztalati vektorok  $\{\mu_{i1}, \mu_{i2}, \dots, \mu_{iM}, i = 1, \dots, N\}$ . Minden állapotra egy halmaz  $M$  prototípus vektor, a jel terének centroidjainak modellezésére minden állapotban.
- Állapot tapasztalati vektor valószínűségi modellje. Lehet diszkrét az  $M$  prototípusvektorokból, azok hozzárendelt sűrűségfüggvényével  $\mathbf{P} = \{P_{ij}(\cdot); i=1, \dots, N, j=1, \dots, M\}$ , vagy folytonos (általában normális)  $\mathbf{F} = \{f_{ij}(\cdot); i=1, \dots, N, j=1, \dots, M\}$  eloszlásfüggvény modell.
- Kezdőállapot valószínűségi vektor  $\boldsymbol{\pi} = [\pi_1, \pi_2, \dots, \pi_N]$

2.2.2.3. A rejtett Markov modell betanítása

Az első lépés, hogy adatbázist alakítsunk ki a modell számára, nagy számú minta vételezésével a vizsgálat tárgyát képező jelből. Tegyük fel, hogy az adatbázis  $L$  vektor értékekkel bíró szekvenciából áll  $[\mathbf{X}]=[\mathbf{X}_k; k=0, \dots, L-1]$ , és minden szekvenciában  $\mathbf{X}_k=[x(t); t=0, \dots, T_k-1]$  változó számú  $T_k$  vektor van. A paraméterek megfelelő beállításához a modell statisztikájának adathalmazát kell előállítani. Valószínűségi értelemben a modell alkalmassága a modell utólagos valószínűségével  $P_{M|X}(M|\mathbf{X})$  adható meg,  $M$  modell  $\mathbf{X}$  adathalmazára. Ez a betanító folyamat az  $M$  modell utólagos valószínűségének és az  $[\mathbf{X}]$  adathalmaznak a maximalizálását célozza, a Bayes szabály alkalmazásával:  $P_{M|X}(M|\mathbf{X}) = \frac{1}{f_X(\mathbf{X})} f_{X|M}(\mathbf{X}|M)P_M(M)$ ,

ahol a nevezőben álló  $f_X(\mathbf{X})$  egy normalizáló hatással bír, és  $P_M(M)$  az  $M$  modell korábbi valószínűsége. Az iménti egyenlet maximalizálása egyenértékű a valószínűségi függvény maximalizálásával, mely a következő egyenlet segítségével fejezhető ki  $\mathbf{X}$  vektorszekvenciára és  $M$  modellre:

$$f_{X|M}(\mathbf{X}|M) = \sum_s f_{X|S,M}(\mathbf{X}|s, M)P_{s|M}(s|M),$$

ahol  $f_{X|S,M}(\mathbf{X}(t)|s(t), M)$  jelöli az eloszlásfüggvényt az  $M$  modell  $\mathbf{X}$  szekvenciájának  $\mathbf{s}=[s(0), s(1), \dots, s(T-1)]$  állapotában, mely a következőképpen fejezhető ki:

$$f_{X|S,M}(\mathbf{X}|s, M) = f_{X|S}(\mathbf{x}(0)|s(0))f_{X|S}(\mathbf{x}(1)|s(1))\cdots f_{X|S}(\mathbf{x}(T-1)|s(T-1)),$$

ahol az  $N$  állapot egyike lehet  $s(t)$   $t$  időpillanatban,  $f_{X|S}(\mathbf{x}(t)|s(t))$  pedig

$f_{X|S,M}(\mathbf{X}(t)|s(t), M)$  rövidített leírása az  $M$  modell  $\mathbf{x}(t)$  eloszlásfüggvényének  $s(t)$

állapotban. A  $\mathbf{s}$  állapotszekvencia Markov valószínűsége a következőképpen adható

$$\text{meg: } P_{s|M}(s|M) = \pi_{s(0)}a_{s(0)s(1)}a_{s(1)s(2)}\cdots a_{s(T-2)s(T-1)}.$$

Az előző három egyenletből az utolsó kettőt behelyettesítve az elsőbe kapjuk a következőt:

$$\begin{aligned} f_{X|M}(\mathbf{X}|M) &= \sum_s f_{X|S,M}(\mathbf{X}|s, M)P_{s|M}(s|M) = \\ &= \sum_s \pi_{s(0)}f_{X|S}(\mathbf{x}(0)|s(0))a_{s(0)s(1)}f_{X|S}(\mathbf{x}(1)|s(1))\cdots a_{s(T-2)s(T-1)}f_{X|S}(\mathbf{x}(T-1)|s(T-1)) \end{aligned}$$

ahol az összegzés végighalad a összes  $\mathbf{s}$  állapotszekvenciát. A betanítási folyamatban az átmeneti valószínűségeket és a tapasztalati eloszlásfüggvények paramétereit az úgy kell kiszámítani, hogy iménti egyenletből a modell valószínűsége maximális legyen. A közvetlen maximalizálás a modell paramétereivel összefüggésben nem triviális megoldás. Következésképpen az egyenlet számításigénye egy megfigyelési

---

szekvencia  $T$  darab vektorára  $O(N^T)$  lesz. Ez még kis számok esetében is nagyon nagy számításigényt jelent. Az állapot–idő diagram ismétlődő struktúrái azt mutatják, hogy sok az ismétlődő számítás az egyenletben, melyet egy hatékony megvalósítással el tudunk kerülni.

### 3. Zajsűrési módszerek

#### 3.1. Zajzár

Érdekes, bár nem túl hatékony megoldás a jelszint folyamatos vizsgálata, és ha az elér egy kellően alacsony értéket, akkor a jelet teljesen kikapcsolva, csendet kapunk eredményül. Ezzel nem csak a zajt állítjuk meg, hanem az alacsony jelszintű hasznos jelet is, és a jel zajterhelése továbbra is megmarad, ha már egy adott szint felett van az értéke. Ez a módszer megoldást jelenthet olyan esetekben, mikor egy berendezés öngerjesztő folyamatának szeretnénk gátat szabni. Ilyen előfordulhat például mikrofonok esetén, melyek egy légtérben vannak az erősített jelüket megszólaltató hangsugárzókkal, és elhelyezéssel már nem tudunk a problémán segíteni (nagy hangerőre van szükség, és bárhová helyezük a mikrofont, egy zárt térben, beindul az oszcilláció). Másik szemléletes példa az elektromos gitárok torzító berendezése, melynek alapvető működéséhez tartozik az átlagosnál nagyobb erősítés. Ilyenkor az érkező jel kikapcsolása a húrok öngerjesztő rezgésének megakadályozását szolgálhatja. Egy ilyen berendezés megfelelő beállítása nehéz, főleg, ha több vagy többféle hangforrás is van, ráadásul a valódi problémára nem jelent megoldást.

#### 3.2. Analóg valós idejű zajsűrők

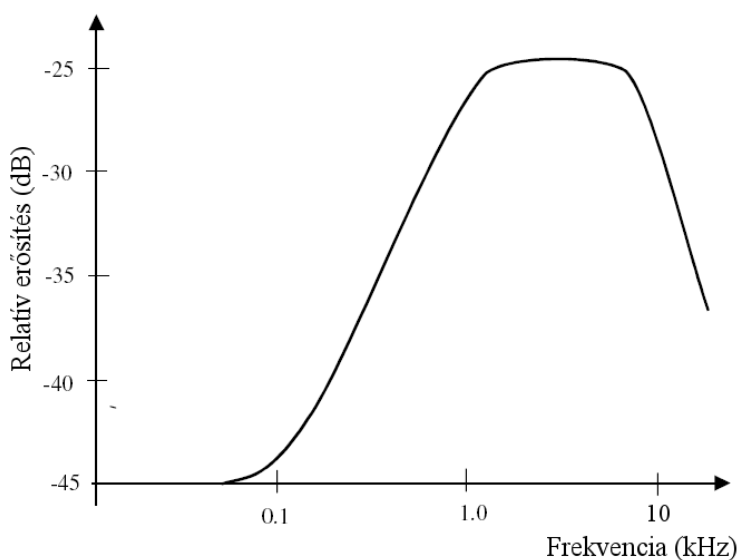
##### 3.2.1. Dolby zajsűrő

A Dolby zajsűrő rendszerek úgy működnek, hogy emelik a magasabb frekvenciákon a hasznos jel amplitúdóját, ezzel javítva a jel–zaj arányt ezen frekvenciákon. A jel nagy része főként az alacsonyabb frekvenciákra – 2 kHz alá – koncentrálódik. A magasabb frekvenciájú jelek, melyek a minőséget és az élményt hordozzák viszonylag alacsonyabb energiájúak, és tovább degradálódnak a zajjal. Például egy mágneses szalagra rögzített jelsorozat esetében a szalag zaja és a korrekciós előerősítő fokozat zaja jelentős magasabb frekvenciájú „sistergést” állít elő. Lejátszásakor ezért ezeken a frekvenciákon romlik a jel–zaj arány – tehát nem lesz egyenletes a teljes átvitt sávban, vagyis az alacsonyabb frekvenciákon jobb, a magasabbakon rosszabb értéket kapunk. Tehát a magasabb frekvenciákon kevesebb hallható hasznos jelet kapunk. A Dolby zajsűrő tehát azon az alapelven működik, hogy a viszonylag kisebb energiájú magasabb frekvenciájú jeleket kiemeli, erősíti a hanganyag felvétele során, a többi jelet pedig változatlanul hagyja. Felvétel közben a jel



egy kiemelés és egy dinamika kompresszió keverékén megy keresztül, melyet lejátszáskor vissza kell alakítani az eredeti állapotába. A kódoló és a dekódoló rendszernek jól illesztettnek kell lennie, a torzítások elkerülése érdekében.

A Dolby laboratóriumban többféle ilyen szűrőt fejlesztettek ki, melyek a Dolby A, Dolby B, és Dolby C. Ezek alapjában véve a sávok számában és az alkalmazott kiemelő stratégiájában különböznek egymástól. A Dolby A a professzionális felhasználás számára lett kifejlesztve, és négy nagy részre osztja a frekvenciaspektrumot – az első sáv 0–80 Hz (alul áteresztő szűrő), a második sáv 80–3000 Hz (hullámsávszűrő), a harmadik sáv 3 kHz körüli felül áteresztő szűrő, a negyedik sáv pedig egy 9 kHz körüli felül áteresztő szűrő segítségével jött létre. A kódoló minden hullámsáv erősítését alkalmazkodóan igazította, az alacsony energiájú jelek kiemelése érdekében. A Dolby A rendszer gondoskodik a 10–15 dB-es erősítésről minden sávban, ha a jel 45 dB-nél nagyobb értékkel esik a maximális rögzítési jelszint alá. A kereskedelmi forgalomba kerülő készülékekben a Dolby B és a Dolby C zajsűrőket alkalmazták. Ezek a rendszerek az imént ismertetett Dolby A rendszerrel szemben már csak két hullámsávon dolgoznak, a négy helyett. A



17. ábra A Dolby C rendszer elő-kiemelő sávátvitele

Dolby B rendszer 10 dB-es emelést biztosít alacsony jelszint esetén (nagyobb, mint 45 dB eltérés a maximális rögzítési jelszinthez képest), a Dolby C rendszer pedig 20 dB-es emeléssel dolgozik, ahogy az a következő ábrán látható.

Ez a módszer csak analóg jelek esetében egy jól meghatározott zajforma csökkentésére alkalmas,

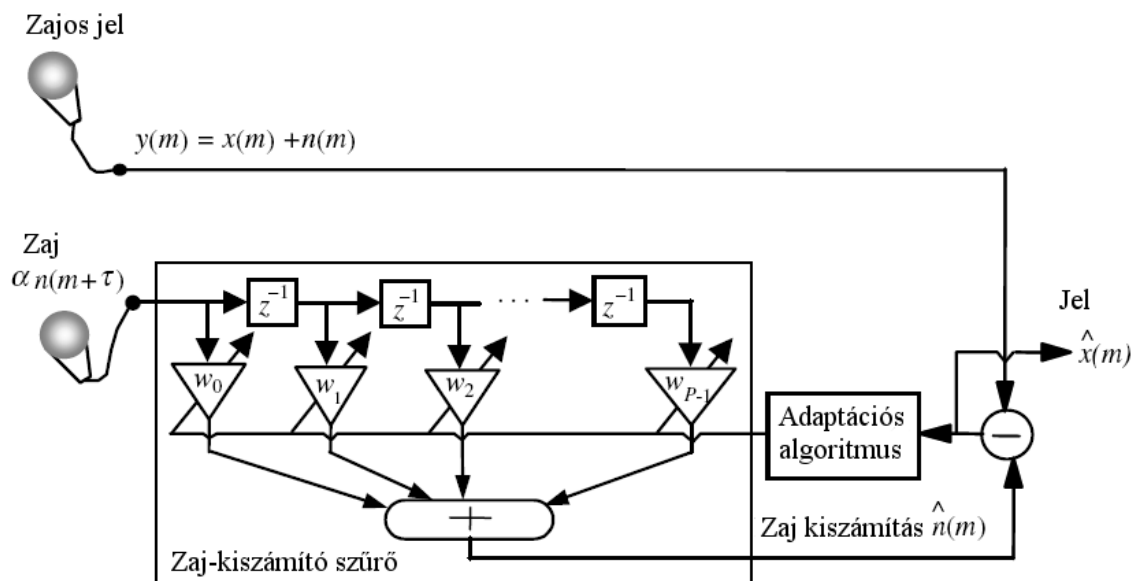
ezért ez nagyon behatárolja az alkalmazási területet. Ezért alkalmazzák ezt a módszert szalagos rögzítő eszközök – úgy mint kompakt kazettás magnetofonok – esetében. Az ilyen készülékek esetében jellemző, hogy a szalagon található mágneses részecskék rendezettsége és az előmágnesezés milyensége befolyásolja a hangminőséget, továbbá a magasabb frekvenciájú hangjelek – az alkalmazott mágne-

szehető anyag függvényében – kisebb jelszinten érkeznek vissza a szalagról, mely egy korrekciós előerősítő fokozat alkalmazását teszi szükségessé, ezáltal az elektronikus rendszernek is lesz egy zaja, melyek együttesének „szűrésére” alkalmazható a Dolby A, B és C rendszer.

### 3.2.2. Zajérvénytelenítés

Egy zajos környezetből (mint például egy mozgó autó) származó beszéd esetén a hasznos jelet a környezet zajával együtt észleljük. A mérőrendszerek esetében az információhordozó jel gyakran a berendezést körülvevő környezet zajával szennyezett. A zajjal együtt észlelt  $y(m)$  jel a következőképpen modellezhető:  $y(m) = x(m) + n(m)$ , ahol  $x(m)$  jelenti a hasznos jelet,  $n(m)$  jelenti a zajt,  $m$  pedig a diszkrét idő mérőszámot. Néhány esetben, például mikor mobiltelefont használunk egy mozgó autóban, vagy rádiókommunikációs eszközt egy repülőgép pilótafülkéjében, talán lehetséges, hogy kimérhető és kiszámítható a környezeti zaj pillanatnyi amplitúdója, egy irányított mikrofon használatával. Ezután a az  $x(m)$  hasznos jel visszaállítható lehet a zajos jel és a zaj különbségeként.

Az adaptív zaj-érvénytelenítő rendszer – két bemenettel rendelkezik, melyekből az egyik a környezeti zajt érzékeli, a másik pedig a zajjal terhelt hasznos jelet – alkalmazható beszédhang kiemelésre. Ebben a rendszerben az egyik irányított mikrofon a zajos  $x(m) + n(m)$  jelet érzékeli, míg a másik irányított mikrofon az előzőtől egy adott távolságra elhelyezve a zajt  $\alpha \cdot n(m + \tau)$  méri. Az a csillapítási szor-



18. ábra Adaptív zajérvénytelenítő elvi felépítése

zótényező és a  $\tau$  időkéleltetési érték meglehetősen túlegyszerűsített modellt ad a mikrofon elhelyezésétől eltérő helyről származó és eltérő zajok hatásainak modellezésére. A második mikrofonból érkező zajt egy adaptív digitális szűrővel dolgozzuk fel, annak érdekében, hogy amplitúdója megegyezzen az első mikrofonban érzékelt hang zajszennyezésének amplitúdójával, ezután pedig ezt a jelet kivonjuk zajos jelből, a zaj érvénytelenítése céljából. Ez a berendezés a zaj alacsonyabb frekvenciáin működik hatékonyan, de szenved a jelek nem állandó jellegétől, és a túlegyszerűsített feltevéstől mely szerint lineáris szűrővel modellezhető a zaj szóródása és terjedése a térben. A berendezés felépítését a 18. ábra szemlélteti.

Sok esetben, például egy telekommunikációs rendszerben alkalmazott vevőkészülék esetében nem ismerjük a zajszennyezés pillanatnyi értékét, csak a zajos jel adott. Ezen esetekben nem érvényteleníthető a zaj, csak mértéke csökkenthető, a jel és a zaj statisztikáját felhasználva. A Wiener szűrő alkalmas ilyen jellegű zajcsökkentési eljárásra, ahol csak a zajos jel adott.

### **3.3. Digitális jelek feldolgoása**

#### *3.3.1. Átalakítások*

##### 3.3.1.1. Analóg–digitális átalakítás

Az analóg jelfeldolgozás során valamilyen folytonosan változó fizikai mennyiséget alakítunk elektronikusan feldolgozható jellé. Ilyen átalakítás történik például egy mikrofonban, ahol a levegő rezgéseit, nyomásváltozásait alakítjuk elektromos feszültséggé, melyet a későbbiekben elektronikus jelerősítő berendezésekkel tudunk tovább feldolgozni. Ezek a jelek még elég sérülékenyek, könnyen terhelődhetnek zajjal, és ezen zajok eltávolítása az analóg jelekből nagyon nehéz feladat. Ha a jelet szeretnénk megszabadítani az additív zajoktól, akkor az első lépés, amit meg kell tennünk a jel átalakítása digitális jellé. A digitális jel fontos jellemzője, hogy egy előre meghatározott értékhalmból vehet fel csak értékeket. Ezen jelek esetében a feldolgozás már matematikai modellek alapján digitális számítógépek segítségével történhet. Továbbá nagyon előnyös tulajdonságaik közé tartozik, hogy sokkal kevésbé érzékenyek a külső additív zajokra, a diszkrét értékek miatt. A legegyszerűbb ilyen digitális jelek kétértékűek, vagy más néven bináris jelek. A mai korszerű kommunikációs rendszerekben gyakran alkalmaznak olyan készülékeket, melyek az analóg jelet először csak digitálissá alakítják, majd ez a digitális jelet

alakítják tovább egy bitfolyammá, melynek értelmezése a vevő oldalon történik. Ez a bit soros adatátvitel, napjaink egyik legkedveltebb átviteli formája.

Az analóg–digitális átalakítás vagy konverzió (Analog to Digital Conversion – ADC) úgy történik, hogy egy analóg feszültségjelet alapul véve abból adott időközönként mintát veszünk. Ez a digitális jel egyik fontos jellemzője, a *mintavételezési frekvencia*. A legjobb minőségű digitális jelet akkor kapnánk, ha egy időegység alatt végtelen sokszor vennénk mintát az analóg jelből, de ez technikailag kivitelezhetetlen és értelmetlen is. A mintavételezés a gyakorlatban rövid időközönként történik:

$T_s = \frac{1}{F_s}$ , ahol  $T_s$  az időközt,  $F_s$  pedig a frekvenciát jelöli. Ez a *mintavételezési eljárás*

az  $x(t)$  analóg jelet egy  $x(m)$  impulzussorozattá alakítja.

Az átalakítás következő lépése a kvantálás, amikor a mintavételezés után kapott impulzusokat egy diszkrét értékhalmazra képezzük. A halmaz elemeit egy analóg jeltartomány meghatározott számú részekre osztásával kapjuk meg. A kvantálás során a mintavételezés eredményeképpen kapott impulzussorozat eleméről eldöntjük, hogy melyik két érték közé (melyik tartományba) esik, és előállítjuk a  $x(m)$  jelből az  $y(m)$  jelet, mely a kvantálás eredményeképpen keletkezik. Ez a digitális jel másik fő jellemzője, amit *kvantálási hossz*nak mondunk. A jel pontossága és a visszaalakításkor kapott eredmény nagymértékben függ a tartományok számától, így a kvantálási hosszról.

A következő lépés a kódolási művelet. Ebben a műveletben minden mintavételi pontban felvett jelhez a kvantálás végeredményének megfelelő bináris kódszót rendelünk hozzá. Ezzel az eljárással minden  $y(m)$  jelhez bináris kódszót rendeltünk, tehát végeredményként kaptunk egy digitálisan feldolgozható jelet.

### 3.3.1.2. Digitális–analóg átalakítás

A digitális jel analóg jellé történő alakításakor (Digital to Analog Conversion – DAC) először a vett bináris jelsorozatot kell helyreállítani, amennyiben az szükséges. A jelsorozat visszaállítását követően a kódszavakból előállítható az  $y'(m)$  függvény, melyet egy alul-áteresztő szűrővel szűrve megkapható az eredeti  $x(t)$  jelet közelítő  $x'(t)$  analóg jel.

A közelítés javítására alkalmazhatunk interpolációt, mely a két függvény hasonlóságát növeli és csökkenti a torzítást. Az analóg jel mintavételezett formájából történő visszaállítás feltétele a Nyquist elméletben megadott. A Nyquist elmélet azt mondja ki, hogy egy sávkorlátozott jel esetében annak  $F_c$  legnagyobb frekvenci-

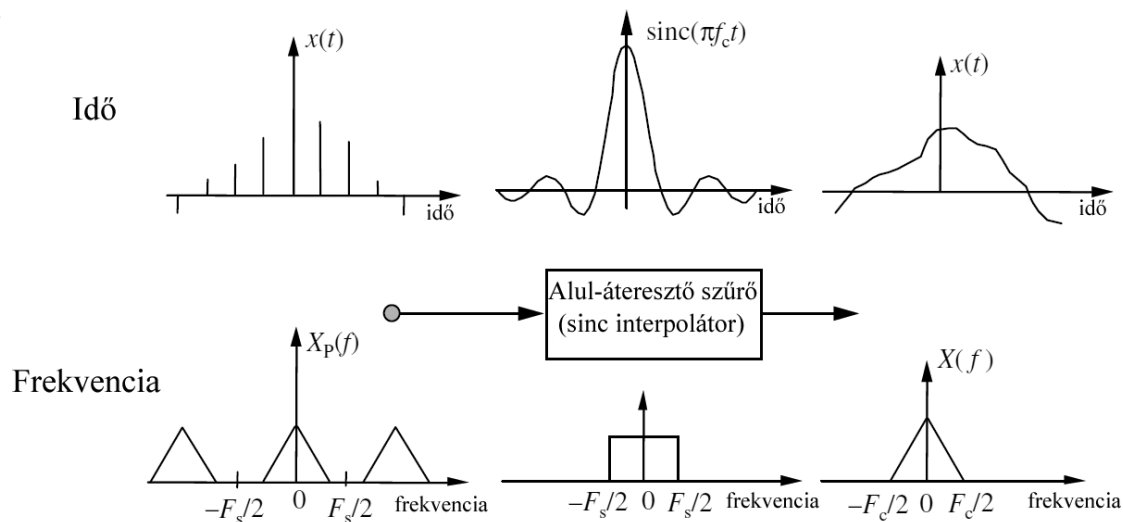
ájú összetevője is visszaállítható abban az esetben, ha a mintavételezés másodpercenkénti frekvenciája nagyobb volt a frekvencia kétszeresénél, vagyis  $F_s > 2F_c$ . Tegyük fel, hogy egy sávkorlátozott  $x(t)$  folytonos jel  $F_s$  másodpercenkénti mintavételezési frekvenciával mintavételezett. A diszkrét időközönként mintavételezett jel  $x(m)$  a következő szorzattal fejezhető ki:  $x(m) = x(t)p(t) = \sum_{m=-\infty}^{\infty} x(t)\delta(t - mT_s)$ , ahol

$p(t) = \delta(t - mT_s)$  a mintavételező függvény, és  $T_s = \frac{1}{F_s}$  a mintavételezési idő intervallum hossza. Vegyük az iménti egyenletet Fourier transzformálva, mely megmutatja, hogy a mintavételezett jel spektruma a következő egyenlettel írható fel:

$X_s(f) = X(f) * P(f) = \sum_{k=-\infty}^{\infty} X(f + kF_s)$ , ahol  $X(f)$  az  $x(t)$  jel spektruma, és  $P(f)$  a  $p(t)$

mintavételező függvénye, a  $*$  pedig a konvolúciós szorzás műveletét jelöli. Az iménti egyenlet állítása szerint a mintavételezett jel spektruma az alapsáv  $X(f)$  spektrumából és annak egyenletes  $T_s = \frac{1}{F_s}$  közönkénti ismétléséből vagy leképezéséből

áll össze. Ha a mintavételezési frekvencia felette van a Nyquist aránynak, akkor az alapsáv  $X(f)$  spektruma nincs átfedésben a  $X(f \pm kF_s)$  képeivel, és az eredeti jel



19. ábra Folytonos jel visszaállítása mintáiból. A frekvenciatartomány interpolációja az alul-áteresztő szűrővel ekvivalens

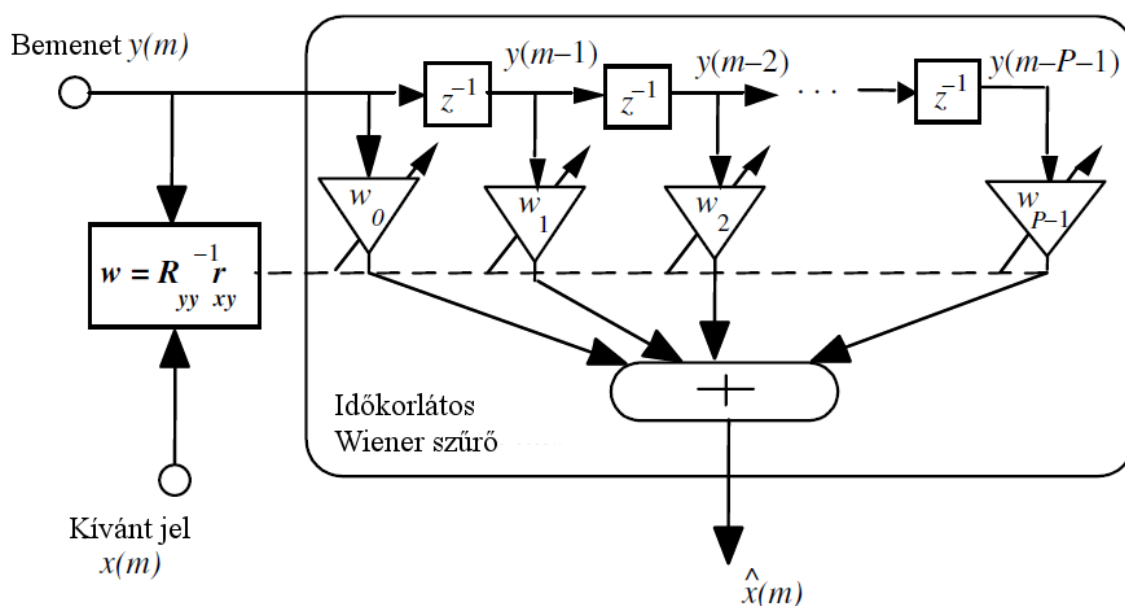
A gyakorlatban az  $F_s$  mintavételezési frekvencia nagyobb mint  $2F_s$ , általában  $2,5F_s$ , annak érdekében, hogy az átmeneti sáv szélességet hozzáigazítsuk az interpoláció alul-áteresztő szűrőjéhez.

### 3.3.2. A Wiener szűrő

A Wiener elméletet kidolgozójáról, Norbert Wiener-ről nevezték el, aki megfogalmazta az adatfüggő lineáris legkisebb négyzetek módszerére épülő szűrő elméletét. A Wiener szűrő központi szerepet játszik alkalmazásainak széles körében, mint például a visszhang érvénytelenítés, jelvisszaállítás, csatorna kompenzáció és rendszerazonosítás. A szűrő együtthatóit úgy kell kiszámítani, hogy a szűrő kimenő jele a lehető legjobban hasonlítson – vagyis különbségük négyzetének várható értéke legyen minimális – egy ismert úgynevezett megkívánt jelre. A Wiener elmélet feltételezi, hogy a jelek állandóak. Azonban ha a szűrő együtthatóit időnként újra kiszámítjuk, minden  $N$  elemű jelblokk után, a szűrő alakítja önmagát, a blokkokon belüli átlagos jellemzőkhöz, így blokk-alkalmazkodóvá válik. A blokk-alkalmazkodó (vagy szegmens-alkalmazkodó) szűrő alkalmazható olyan jelek esetében, mint például az emberi beszéd.

#### 3.3.2.1. A legkisebb négyzetek becslése

Wiener eredeti munkájában a folyamatos idejű legkisebb négyzetek módszerét adta meg, interpolációval. A módszert később kiterjesztette diszkrét időre is, mely egyszerűen megoldható, és a gyakorlatban jobban hasznosítható a digitális jelpro-



20. ábra Wiener szűrő felépítése

cesszorokban. A szűrő lehet végtelen hosszú jel feldolgozását célzó, vagy véges jel-hosszúságot figyelembevevő. Általában a végtelen szűrőnek egy halmaz nem lineáris egyenlet az eredménye, a véges hosszúságú jelre alkalmazott szűrőnek pedig egy halmaz lineáris egyenlet, mely egy zárt alakú megoldás. A véges szűrő viszonylag egyszerűbb számítást tesz lehetővé, működéséből adódóan stabilabb és praktikusabb is. Hátránya a végtelen szűrővel szemben, hogy a kívánt eredmény elérése érdekében nagyon sok együttható szükséges.

A 20. ábrán szemléltetett Wiener szűrő egy  $\mathbf{w}$  együttható vektorral megvalósított. A szűrő kap egy  $y(m)$  bemeneti jelet, és készít egy  $\hat{x}(m)$  kimeneti jelet, mely a legkisebb négyzetes eltérést mutatja az  $x(m)$  megkívánt vagy céljelhez képest. A szűrő bemenete és kimenete közötti összefüggés a következő egyenlettel írható fel:

$$\hat{x}(m) = \sum_{k=0}^{P-1} w_k y(m-k) = \mathbf{w}^T \mathbf{y}, \text{ ahol } m \text{ a diszkrét idő mérőszám, a szűrő bemeneti}$$

jele:  $\mathbf{y}^T = [y(m), y(m-1), \dots, y(m-P+1)]$ , és a paramétervektor (a Wiener szűrő együttható vektora)  $\mathbf{w}^T = [w_0, w_1, \dots, w_{P-1}]$ . A szűrési művelet kétféle egymással ekvivalens alakban fejthető ki, konvolúciós összegként, és a belső vektorok szorzataként. A hibajel  $e(m)$  a kimenőjel  $\hat{x}(m)$  és a megkívánt jel  $x(m)$  különbségeként számítható ki a következő összefüggéssel:  $e(m) = x(m) - \hat{x}(m) = x(m) - \mathbf{w}^T \mathbf{y}$ . Ebből az egyenletből kiderül, hogy egy adott  $y(m)$  bemenőjelre és  $x(m)$  megkívánt jelre a hibajel a  $\mathbf{w}$  együtthatóvektortól függ. A hibajel és az együtthatóvektor közötti összefüggés kiderítése érdekében kifejtve az előző egyenletet a következőt kapjuk:

$$\begin{pmatrix} e(0) \\ e(1) \\ e(2) \\ \vdots \\ e(N-1) \end{pmatrix} = \begin{pmatrix} x(0) \\ x(1) \\ x(2) \\ \vdots \\ x(N-1) \end{pmatrix} - \begin{pmatrix} y(0) & y(-1) & y(-2) & \dots & y(1-P) \\ y(1) & y(0) & y(-1) & \dots & y(2-P) \\ y(2) & y(1) & y(0) & \dots & y(3-P) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y(N-1) & y(N-2) & y(N-3) & \dots & y(N-P) \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_{P-1} \end{pmatrix}.$$

Ez az egyenlet  $N$  mintára,  $y(m)$  és  $x(m)$  jelekre lett felírva. Tömörített alakban a következőképpen írható fel:  $\mathbf{e} = \mathbf{x} - \mathbf{Y}\mathbf{w}$ , ahol  $\mathbf{e}$  a hibavektor,  $\mathbf{x}$  a kívánt jel vektor,  $\mathbf{Y}$  a bementi jel mátrix, és  $\mathbf{Y}\mathbf{w}$  szorzat pedig a  $\hat{x}$  kimenőjel vektor. Feltételezzük, hogy  $P$  kiinduló minta ismert, vagy hanem akkor 0 értékűnek tekintendő. Ha az előbb felírt mátrixegyenletben az  $N$  és a  $P$  értéke egyenlő, akkor egy négyzetes mátrixegyenletet kapunk, és ez az egyetlen megoldás, ahol a számítási hiba értéke nulla. Ha  $N < P$  a minták száma  $N$  elégtelen mennyiségű ahhoz, hogy elérjük az egyedülálló megoldást a szűrő együtthatóinál, ezért ebben az esetben végtelen sok megoldás létezik nulla számítási hibával és a mátrixegyenletet határozatlannak mond-

juk. A gyakorlatban a jelminták száma sokkal nagyobb, mint a szűrő hossza – vagyis  $N > P$  –, ebben az esetben a mátrixegyenletet túlhatározottnak mondjuk, és létezik egy egyedi megoldás, általában nulla számítási hibával. Mikor  $N > P$  a szűrő együtthatói úgy lesznek kiszámítva, hogy az átlagos hiba minimális legyen, mint például a hiba átlagos értékének abszolút értéke  $E[|e(m)|]$ , vagy a hiba négyzetének középértéke  $E[e^2(m)]$ , ahol  $E[.]$  a várható érték operátor. A hibafüggvény kiválasztása hatással van a megoldás optimális mivoltára, és a számítási komplexitására. A Wiener elméletben az objektív kritérium a legkisebb négyzetes hiba (LSE<sup>2</sup>) a szűrő kimeneti jele és a kívánt jel között. A legkisebb négyzetes hiba kritérium optimális a normális eloszlást követő jelek esetében. A korlátozott idő intervallumra alkalmazott legkisebb négyzetes hiba kritériumú Wiener szűrő egy lineáris, zárt formulához vezet. A Wiener szűrő együtthatói egy átlagos négyzetes hiba függvény  $E[e^2(m)]$  minimalizálásával eszközli ki, tekintettel a szűrő  $\mathbf{w}$  együttható vektorára. A számítási hiba négyzetének középértéke a következő egyenlettel adható meg:  $E[e^2(m)] = E[(x(m) - \mathbf{w}^T \mathbf{y})^2] = E[x^2(m)] - 2\mathbf{w}^T E[\mathbf{y}x(m)] + \mathbf{w}^T E[\mathbf{y}\mathbf{y}^T] \mathbf{w} = r_{xx}(0) - 2\mathbf{w}^T \mathbf{r}_{yx} + \mathbf{w}^T \mathbf{R}_{yy} \mathbf{w}$ , ahol  $\mathbf{R}_{yy} = E[\mathbf{y}(m)\mathbf{y}^T(m)]$  a bemenő jel autokorrelációs<sup>3</sup> mátrixa,  $\mathbf{r}_{xy} = E[x(m)\mathbf{y}(m)]$  pedig a bemeneti jel és a kívánt jel kereszt-kölcsönösségi vektora. Az előző kifejezést kifejtve a következő egyenletet kapjuk:

$$E[e^2(m)] = r_{xx}(0) - 2 \sum_{k=0}^{P-1} w_k r_{yx}(k) + \sum_{k=0}^{P-1} w_k \sum_{j=0}^{P-1} w_j r_{yy}(k-j), \text{ ahol } r_{yy}(k) \text{ és } r_{yx}(k) \text{ az } \mathbf{R}_{yy}$$

autokorrelációs mátrix és az  $\mathbf{r}_{xy}$  kereszt-kölcsönösségi vektor elemei külön-külön. Az előző egyenletből a korlátozott idő intervallumra alkalmazott szűrő négyzetes hibájának középérték számítására egy másodfokú függvény alkalmazható a  $\mathbf{w}$  együttható vektor értékeire, mely függvénynek egyetlen minimuma van. Például ha szűrőnek csak két együtthatója van ( $w_1, w_2$ ), akkor a négyzetes hiba középérték függvénye egy medence alakú képet mutat, egyetlen minimummal. A legkisebb négyzetes hiba jelenti az energiaminimumot.

### 3.3.3. Zajcsökkentés spektrális kivonással

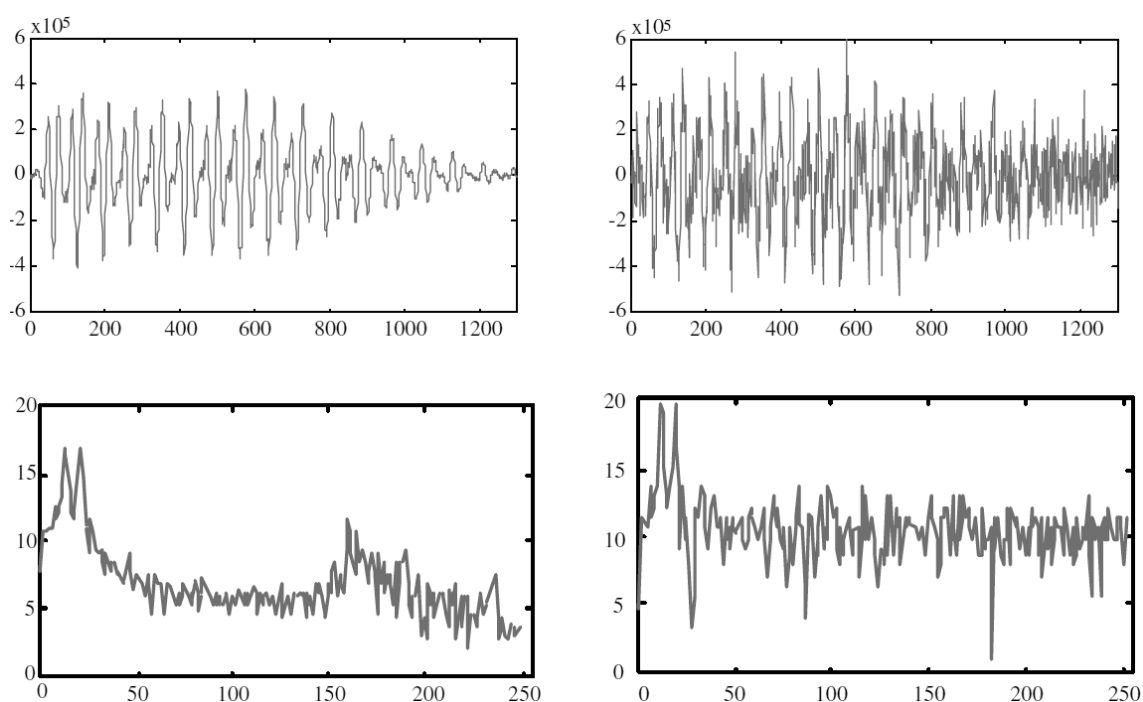
A spektrális kivonás módszerének lényege, hogy a zajos jelből veszünk egy olyan szakaszt, ahol a hasznos jel nincs jelen, tehát csak zaj észlelhető, és ezt a jelet egy megfelelő módszerrel kivonjuk a zajos jelből. Ez megtehető oly módon, hogy a jelek statisztikáit és spektrális paramétereit vesszük alapul. Ezzel a módszerrel adott esetben jelvisszaállítás is lehetséges.



A problémára a legegyszerűbb megoldást jelentené, ha külön rendelkezésre állna a zajos jel és a zaj. Ebben az esetben egy egyszerű művelettel eltávolítható lenne a zaj. A zajérvénytelenítő, melynek működését már korábban kifejtettem pontosan ilyen elven működik. Általában azonban csak a zajos jel áll rendelkezésre, tehát nincs lehetőség a zaj érvénytelenítésére, de előfordulhat, hogy a zaj hatását csökkenteni tudjuk a hasznos jel spektrumán.

### 3.3.3.1. A spektrális kivonás módszere

A 21. ábra olyan jeleket mutat, melyek zajjal terhelték, de nincs lehetőség a zaj érvénytelenítésére. A hasznos jel spektrumának középvértéke növelhető úgy, hogy a

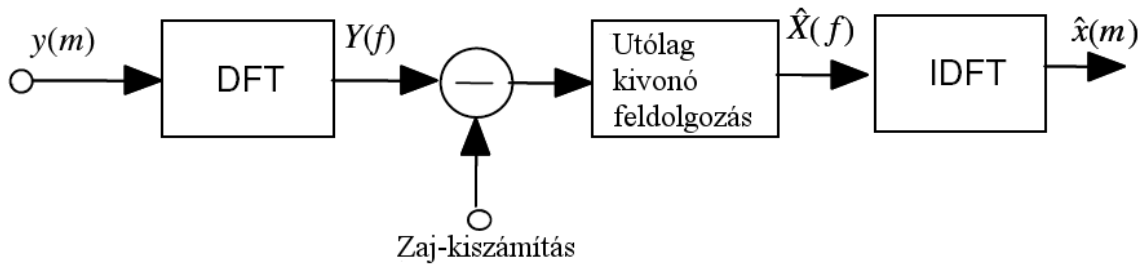


21. ábra Zajjal terhelt jelek alakjai az idő függvényében

zajos jel spektrumának középvértékéből kivonjuk a zaj spektrumának középvértékét. A zajos jel időben a következőképpen modellezhető:  $y(m)=x(m)+n(m)$ , ahol  $y(m)$  jelöli a zajjal terhelt jelet,  $x(m)$  a hasznos jelet,  $n(m)$  a zajt,  $m$  pedig a diszkrét idő mérőszám. Egy frekvenciatartományban az előző egyenlet a következőképpen fejthető ki:  $Y(f)=X(f)+N(f)$ , ahol  $Y(f)$  az  $y(m)$  zajos jel Fourier transzformáltja,  $X(f)$  az  $x(m)$  hasznos jel Fourier transzformáltja,  $N(f)$  az  $n(m)$  zaj Fourier transzformáltja, és  $f$  a frekvenciaváltozó. A spektrális kivonásban az  $y(m)$  zajos jelet vesszük, ezt egy átmeneti tárolóba helyezzük, majd felosztjuk  $N$  egyenlő részre. Minden szegmens esetében Hann vagy Hamming ablak segítségével szűrjük a spektrumot, ezután transzformáljuk diszkrét Fourier transzformációval (DFT) az összes  $N$  darab spektrális mintát. Az ablakok minden szegmens végén csökkentik annak hatását,

hogy a vett jelet feldaraboltuk, vagyis nem folytonos jeleket illesztünk össze. Az ablakozott jel így adható meg:  $y_w(m) = w(m)y(m) = w(m)[x(m) + n(m)] = x_w(m) + n_w(m)$ . Az alakozási eljárás a frekvenciatartományokra az alábbi módon fejthető ki (DFT segítségével):  $Y_w(f) = W(f) * Y(f) = X_w(f) + N_w(f)$ , ahol a  $*$  operátor a konvolúciós szorzást jelenti.

A következő ábra a spektrális kivonás módszerének blokkdiagramját szemlélteti. (Az IDFT jelölés a diszkrét Fourier transzformáció inverzét jelenti)



22. ábra A spektrális kivonás egyszerűsített blokkdiagramja

A spektrális kivonás egyenlettel a következő módon írható le:  $|\hat{X}(f)|^b = |Y(f)|^b - \alpha \overline{|N(f)|^b}$ , ahol  $|\hat{X}(f)|^b$  az eredeti jel spektrumának kiszámítása,  $|X(f)|^b$  és  $|N(f)|^b$  pedig az időátlagolt zaj spektrum. A spektrális kivonás nagyságához a  $b$  kitevő értékét 1-re, az erősségéhez pedig 2-re kell választani. Az  $\alpha$  paraméter a kivont zaj mennyiségét szabályozza. A teljes kivonáshoz  $\alpha=1$  értéket kell választani, a túlvonáshoz pedig  $\alpha>1$  érték lesz a megfelelő. Az időátlagolt zajspektrum beszerezhető a jelmentes szakaszokból, ahol csak kizárólag a zaj van jelen, nincs hasznos jel, és az alábbi összefüggés segítségével formalizálható:

$$\overline{|N(f)|^b} = \frac{1}{K} \sum_{i=0}^{K-1} |N_i(f)|^b, \text{ ahol } |N_i(f)|^b \text{ a spektrum } i\text{-edik zajkerete, és feltesszük,}$$

hogy a zaj  $K$  keretből áll, és  $K$  változó. Más esetben az átlagolt zajspektrum beszerezhető az elsőrendű digitális alul-áteresztő szűrő kimeneteként a következőképpen:

$$\overline{|N_i(f)|^b} = \rho \overline{|N_{i-1}(f)|^b} + (1 - \rho) |N_i(f)|^b, \text{ ahol az alul-áteresztő szűrő együtthatója}$$

$\rho$ , melynek tipikus értéke 0,85 és 0,99 közé szokott esni. Az időtartomány alapú jelvisszaállításhoz a nagyságspektrum  $|\hat{X}(f)|$  kiszámítását kombináljuk a zajos jel fázisával, majd ezt követően transzformáljuk időtartományba az inverz diszkrét Fourier transzformáció segítségével, mely a következőképpen írható fel:

$$\hat{x}(m) = \sum_{k=0}^{N-1} |\hat{X}(k)| e^{j\theta_Y(k)} e^{-j\frac{2\pi}{N}km}, \text{ ahol } \theta_Y(k) \text{ a zajos jel fázisa } Y(k) \text{ frekvencián. Ez a jel-}$$

viSSzaállító egyenlet feltételezi, hogy a hallható tartományba eső zaj főként a nagyságspektrum torzítása és a fázistorzítás miatt nagyrészt nem hallható. A szimulált fázistorzítás megfigyelhető hatásának becslése igazolta a feltevést. A zaj spektrumának változatossága következtében előfordulhat, hogy nagyság és az energiaspektrum szempontjából negatív eredményt kapunk a spektrális kivonással. Ennek végeredményeképpen nagyon valószínű, hogy csökkenni fog a jel-zaj arány. A negatív eredmények elkerülése érdekében a spektrális kivonás kimeneti jelét egy utófeldolgozó térképező függvénnyel feldolgozzuk, melyet a következőképpen formalizálhatunk:  $T[|\hat{X}(f)|] = \begin{cases} |\hat{X}(f)| & \text{ha } |\hat{X}(f)| > \beta |Y(f)| \\ \text{fn}[|Y(f)|] & \text{egyébként} \end{cases}$ . Például választhatunk egy

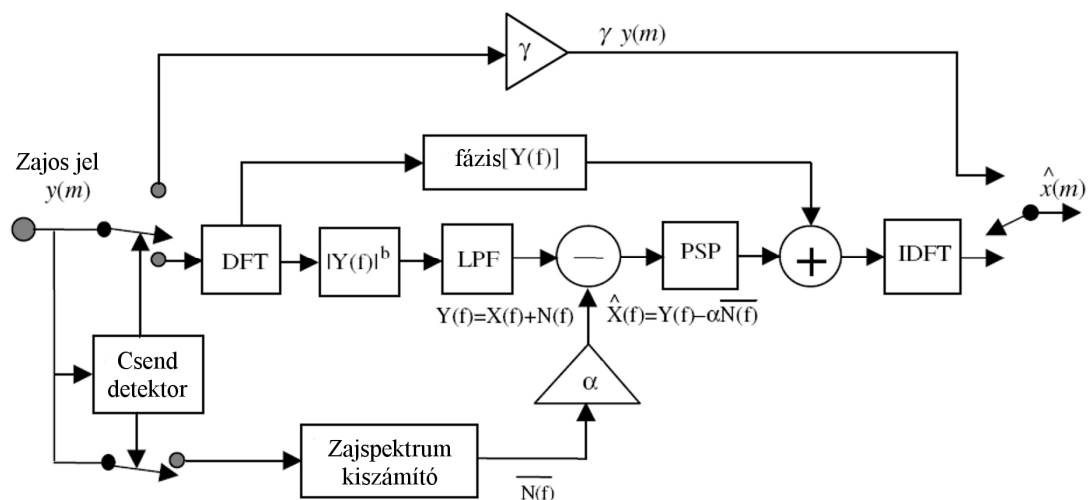
olyan szabályt, mely ha a  $|\hat{X}(f)| > 0,01 |Y(f)|$  (a nagyság spektrum szorzása 0,01-dal ekvivalens a  $-40\text{dB}$ -lel) érték igaz, akkor az  $|\hat{X}(f)|$  érték valamilyen függvénye legyen a zajos jelnek  $\text{fn}[|Y(f)|]$ . Ennek legegyszerűbb formája, hogy beállítunk egy  $\text{fn}[|Y(f)|] = \text{zajszint}$  értéket, ahol a zajszint pozitív konstans. Egy másik lehetőség, ha  $\text{fn}[|Y(f)|] = \beta |Y(f)|$ . Ez esetben:

$$T[|\hat{X}(f)|] = \begin{cases} |\hat{X}(f)| & \text{ha } |\hat{X}(f)| > \beta |Y(f)| \\ \beta |Y(f)| & \text{egyébként} \end{cases}$$

A spektrális kivonás energiaspektrum vagy nagyságspektrum tartományok alapján is implementálható. A két módszer hasonló, de elméletileg eredményeznek némi különbséget a várható teljesítményben.

### 3.3.3.2. A spektrális kivonó megvalósítása

Egy spektrális zajkivonó rendszer blokkdiagramját mutatja az alábbi ábra,



23. ábra Spektrális kivonó megvalósításának blokkdiagramja

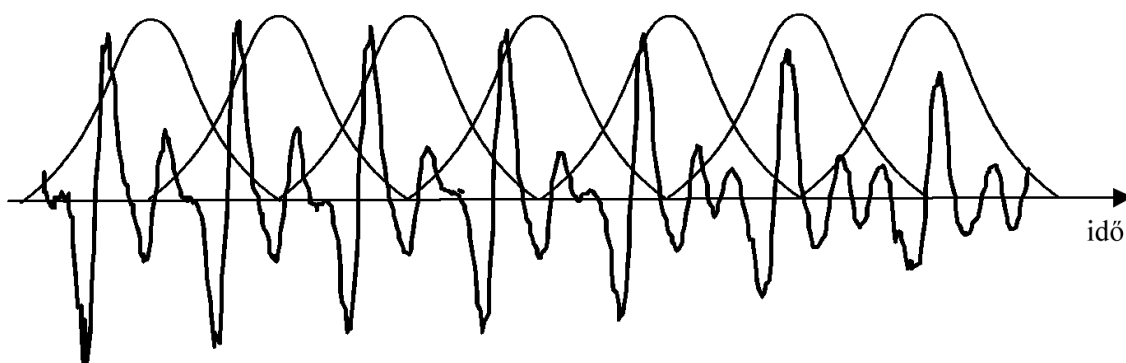
mely a következő alrendszeréből áll:

- a.) egy csend detektor a jelmentes időszakok megtalálására, a zaj spektruma ezen időtartamok alatt frissül.
- b.) a diszkrét Fourier transzformációt megvalósító rész (DFT) az időtartományt frekvenciatartományá alakítja, a DFT egységet egy nagyság operátor követi.
- c.) alul-áteresztő szűrő (Lowpass Filter, LPF), a zaj egyenetlenségeinek csökkentésére, melynek célja a zaj ingadozásaiból adódó feldolgozási torzítás csökkentése.
- d.) utófeldolgozó, a spektrális kivonás eredményeképpen keletkezett torzítások eltávolítására.
- e.) inverz diszkrét Fourier transzformációt végző rész, mely frekvenciatartományból viszi a feldolgozott jelet időtartományba.
- f.) csillapító berendezés  $\gamma$ , mely a zaj csillapítását végzi a csöndes részeken.

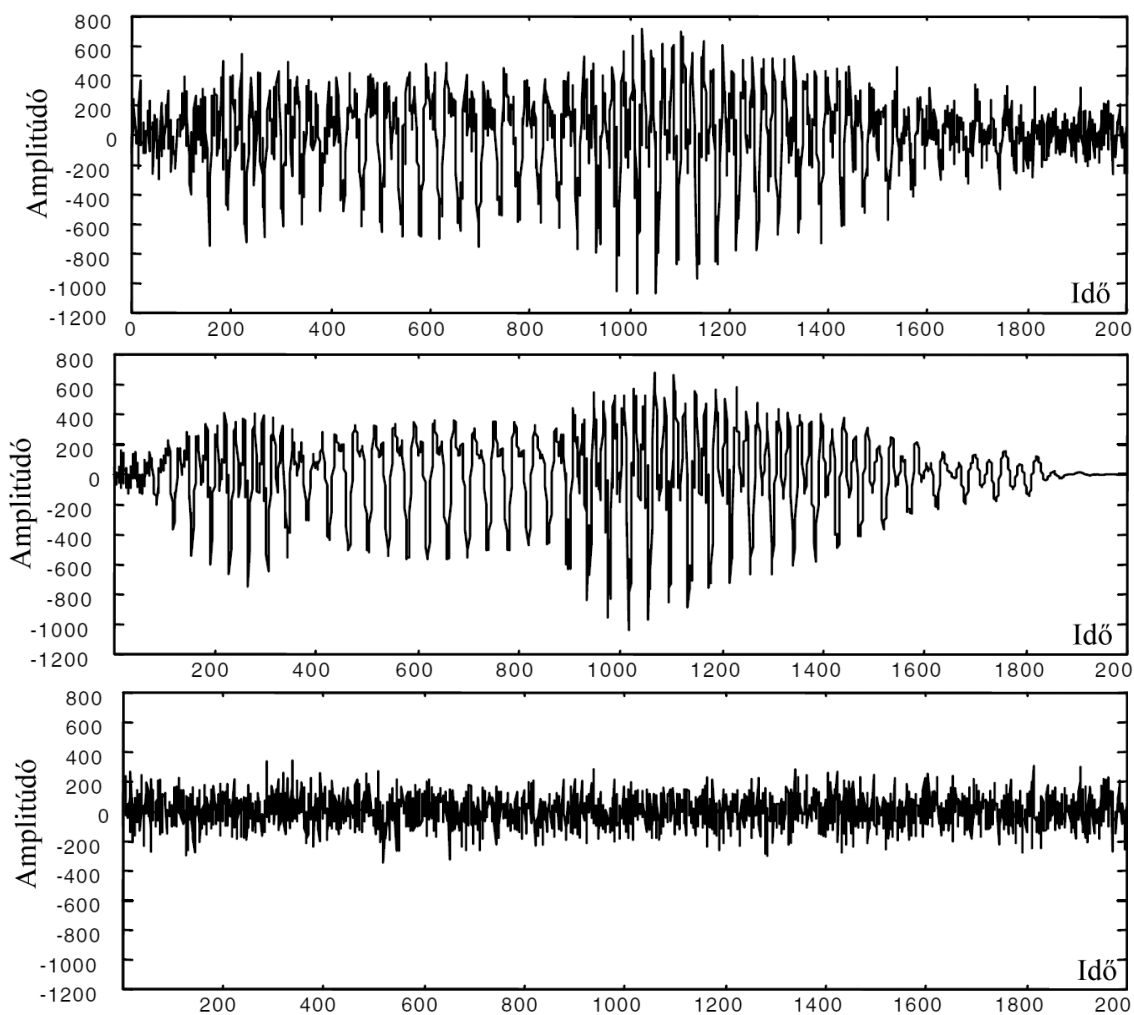
A diszkrét Fourier transzformáción alapuló spektrális kivonó blokkdiagramja egy algoritmus. A beérkező audiojelet átmeneti tárolóba rakjuk, majd  $N$  darab egymást átfedő mintablokkra osztjuk. Minden blokk Hann vagy Hamming ablakolt, és azt követően diszkrét Fourier transzformációval átalakítjuk frekvenciatartományá. A spektrális kivonás után a jel nagyságspektrumát kombináljuk a zajos jel fázisával, és visszaalakítjuk időtartományá. Minden blokkot egymáshoz illesztünk, úgy, hogy legyen átfedés a megelőző és a rákövetkező blokkal, és ez adja a végleges kimeneti jelalakot.

A blokk hosszának megválasztása a spektrális elemzéshez egy kompromiszsum, mely két ellentmondó szükséglet, az időfelbontás és a spektrális felbontás igényeit kell, hogy kielégítse. Tipikusan 5–50 ms közé esik a használt blokkok hossza. Ez egy 20 kHz-es mintavételezési frekvencián azt jelenti, hogy  $N$  értéke 100–1000 tartományba esik. A spektrum frekvenciafelbontása egyenesen arányos a minták számával ( $N$ ). Nagyobb  $N$  esetén a spektrum értékelése jobb lesz. Ez csak részben igaz a frekvenciatartomány alsó részére, mivel az alacsonyabb frekvenciájú részek az idő múlásával lassan változnak, és nagyobb ablak szükséges a stabil értékeléshez. Az ellentmondó szükséglet az audiojelek nem stacionárius természetéből adódnak, tehát az ablak méretének azért nem szabad túl nagyoknak lennie, hogy a rövid időtartamú események ne mosódjanak el.

Az ablakok átfedéseinek (24. ábra) fő feladata, hogy csillapítsák a blokkok végpontjain keletkező folytonossághiányokat, melyek a darabolásból adódnak. Bár

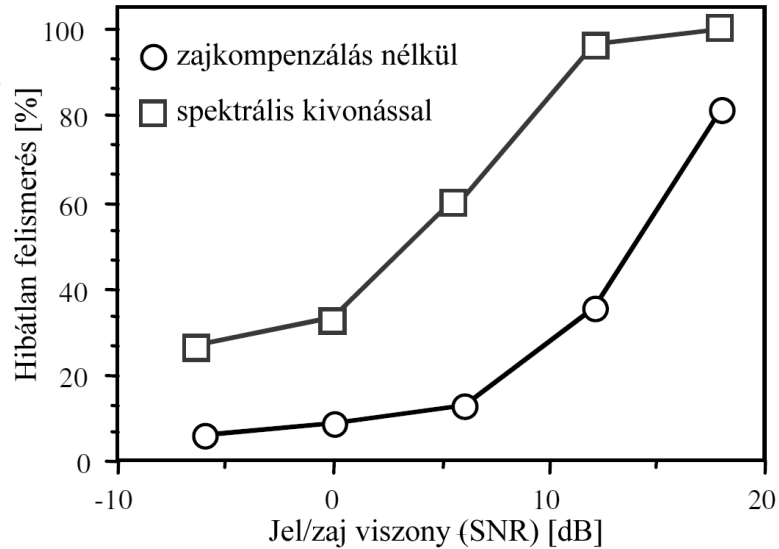


24. ábra A spektrális kivonásban alkalmazott ablakolási és átfedési folyamat létezik számos felhasználható ablaktípus, különböző karakterisztikákkal, mégis a legtöbb megvalósításban Hamming ablakokat használnak. A torzítás eltávolítás során – melyet a spektrális kivonás eredményezett – az utófeldolgozó algoritmus olyan információkat használ fel, mint a frekvenciacsatornák összefüggése az egymást követő blokkokban, valamint az jel eseményeinek és torzításának időtartamai.



25. ábra Zajos jel (felső), zajmentes jel (középső), és zaj (alsó)

A jel spektrális összetevőinek összefüggése – az időtengely mentén – részben szabályozható az ablakhossz megválasztásával és az átfedéssel. Ez az összefüggés növekszik az ablak hosszának csökkentésével és az átfedés növelésével. Az átfedés növelése azonban növeli a zajfrekvenciák összefüggését is, az időtengely mentén.



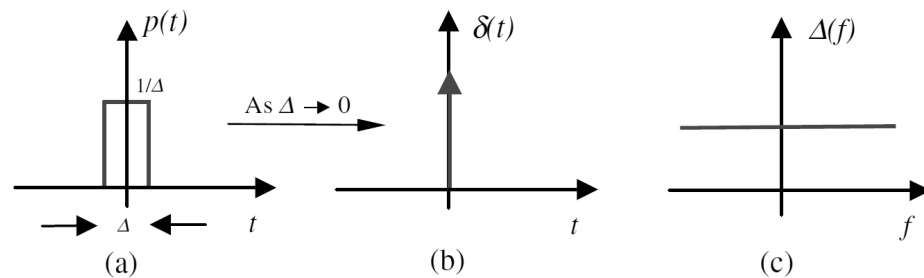
26. ábra A spektrális kivonás jelentősége a beszéd felismerésben

### 3.3.4. Az impulzív zaj eltávolítása

Az impulzív zajok gyors be/ki kapcsolásokhoz hasonló impulzusokat produkálnak és nagyon kiszámíthatatlan a megjelenésük, tehát bármely időpillanatban előfordulhatnak, és bármekkora időközrel. Ezen zajok eltávolítására és csökkentésére az előző részben kifejtett spektrális kivonás módszere nem alkalmas, mert itt a zajszint nem határozható meg.

#### 3.3.4.1. Impulzív zaj matematikai modellje

Az analóg impulzív zajt matematikai értelemben az alábbi ábra segítségével szemléltethetjük. Ha megvizsgáljuk az ábra (a) részén szemléltetett egységnyi te-



27. ábra (a) Egységnyi területű lökés, (b) impulzus, (c) az impulzus spektruma

rületű lökeshullám  $p(t)$  (négyzög jel) viselkedését, ahogy a lökés szélessége  $\Delta$  nullához tart, a lökés impulzussá válik. Az impulzus függvényt az ábra (b) része mutatja, és egy végtelenül kicsi időszélességgel rendelkező lökést a következőképpen

definiálhatunk:  $\delta(t) = \lim_{\Delta \rightarrow 0} p(t) = \begin{cases} \frac{1}{\Delta}, & |t| \leq \frac{\Delta}{2} \\ 0, & |t| > \frac{\Delta}{2} \end{cases}$ . Az impulzusfüggvény integrálja a kö-

vetkező:  $\int_{-\infty}^{\infty} \delta(t) dt = \Delta \times \frac{1}{\Delta} = 1$ . Az impulzusfüggvény Fourier transzformáltja felállítha-

tó:  $\Delta(f) = \int \delta(t) e^{-j2\pi ft} dt = e^0 = 1$ , ahol  $f$  a frekvenciaváltozó. Az impulzusfüggvényt

teszt jelleggel is szokták alkalmazni, a rendszer impulzusreakciójának megállapítására. Az ábra (c) részén mutatottak szerint az impulzus egy spektrálisan nagyon gazdag jel, minden frekvenciát azonos mennyiségben tartalmaz, tehát ezért alkalmas a tesztelésre.

Egy digitális impulzus  $\delta(m)$  (8. oldal 4. ábra (a) részén látható) úgy definiálható, mint egy egy minta hosszúságú „be” jel, és kifejezhető:  $\delta(m) = \begin{cases} 1, & m = 0 \\ 0, & m \neq 0 \end{cases}$ , ahol az

$m$  változó jelöli a diszkrét idő mérőszámot. A Fourier transzformációt használva

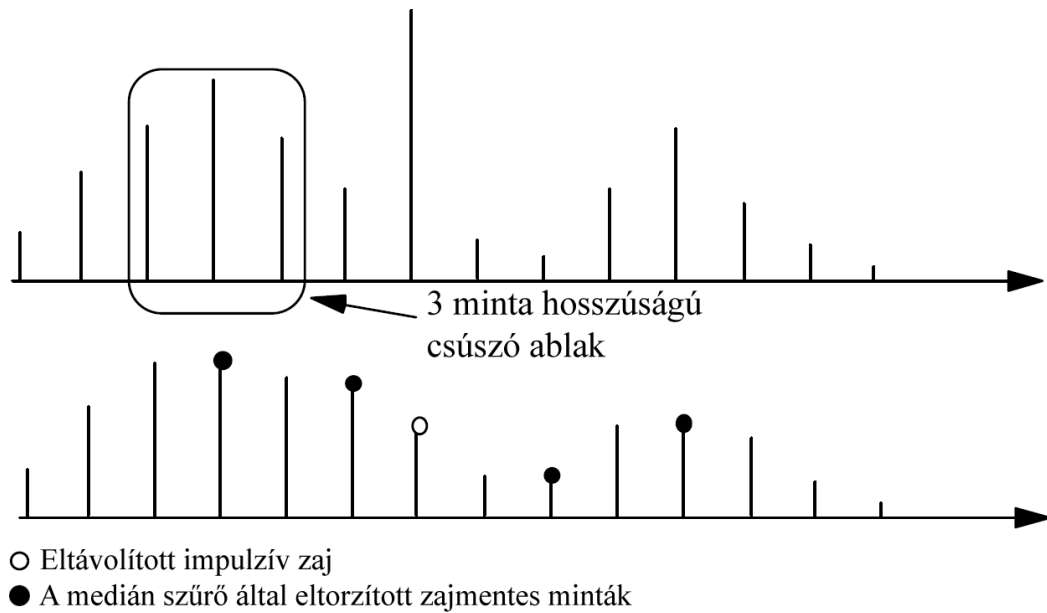
egy digitális impulzus frekvenciaspektruma megadható:  $\Delta(f) = \sum_{m=-\infty}^{\infty} \delta(m) e^{-j2\pi fm} = 1, 0,$

$-\infty < f < \infty$ .

Kommunikációs rendszerekbe a valódi impulzív típusú zajok hossza normálisan nagyobb, mint egy minta.

#### 3.3.4.2. A medián szűrő

Az impulzív zaj eltávolításának klasszikus megközelítése a medián szűrő. Medián alatt értjük az  $\{x(m)\}$  mintahalmaz elemeiből azt az  $x_{\text{med}}(m)$  értéket, melyre igaz, hogy a halmaz elemeinek fele ezen  $x_{\text{med}}(m)$  értéknél nagyobb, a másik fele pedig az  $x_{\text{med}}(m)$  értéknél kisebb. Tekintsük a mediánt minták egy halmazának, melyet úgy kapunk meg, hogy a mintákat rendezzük növekvő vagy csökkenő sorrendbe és kiválasztjuk a középső értéket. A medián szűrés esetén egy előre meghatározott hosszúságú ablakot csúsztatunk folyamatosan a jelen, és az ablakban található középső mintát kicseréljük az ablak mintáinak mediánjára, ahogy azt a következő ábra mutatja.



28. ábra Medián szűrő be- és kimenete. A működés közben torzítás is keletkezhet

A medián szűrő  $\hat{x}(m)$  kimenete az  $y(m)$  bemenet és egy  $2K+1$  mintát tartalmazó hosszúságú medián ablak alapján:

$$\hat{x}(m) = y_{\text{med}}(m) = \text{medián}[y(m-K), \dots, y(m), \dots, y(m+K)]$$

Egy számhalmaz mediánja a halmaz nem lineáris statisztikája egy nagyon jól hasznosítható tulajdonsággal, hogy nem érzékeny egy szokatlanul nagy mintaérték jelenlétére a halmazban, melyet kívülállónak hívunk. Ezzel szemben a középérték és részben a variancia érzékeny egy számhalmazban jelen levő kiugró értékre, vagyis az impulzív típusú zajra. Fontos tulajdonsága a medián szűrőnek – mely részben használható a képfeldolgozásban –, hogy megőrzi az éleket vagy a lépésenkénti folytonossághiányokat a jelben. A medián szűrők hatékonyan használhatók a képfeldolgozásban az impulzív zaj eltüntetésére az élek elmosása nélkül – ez nagyon jelentős a képfeldolgozásban. A medián filterrel történő kísérletezés során kiderült, hogy nem alkalmas magas minőségű eredmények produkálására audiojelek visszaállítása esetében. A medián szűrő nem tud mit kezdeni az olyan impulzív jellegű zajjal, mely hosszabb, mint egy vagy két minta. Továbbá a medián szűrő a hibásan impulzív zajnak vélt jelek módosításával jelentős mennyiségű feldolgozási torzítást okoz. A medián szűrő teljesítménye javítható egy alkalmazkodó határérték bevezetésével, tehát a mintát csak akkor cseréljük ki, ha annak eltérése a középső érték-

től meghaladja a határértéket:  $\hat{x}(m) = \begin{cases} y(m) & \text{ha } |y(m) - y_{\text{med}}(m)| < k\theta(m) \\ y_{\text{med}}(m) & \text{egyébként} \end{cases}$ , ahol

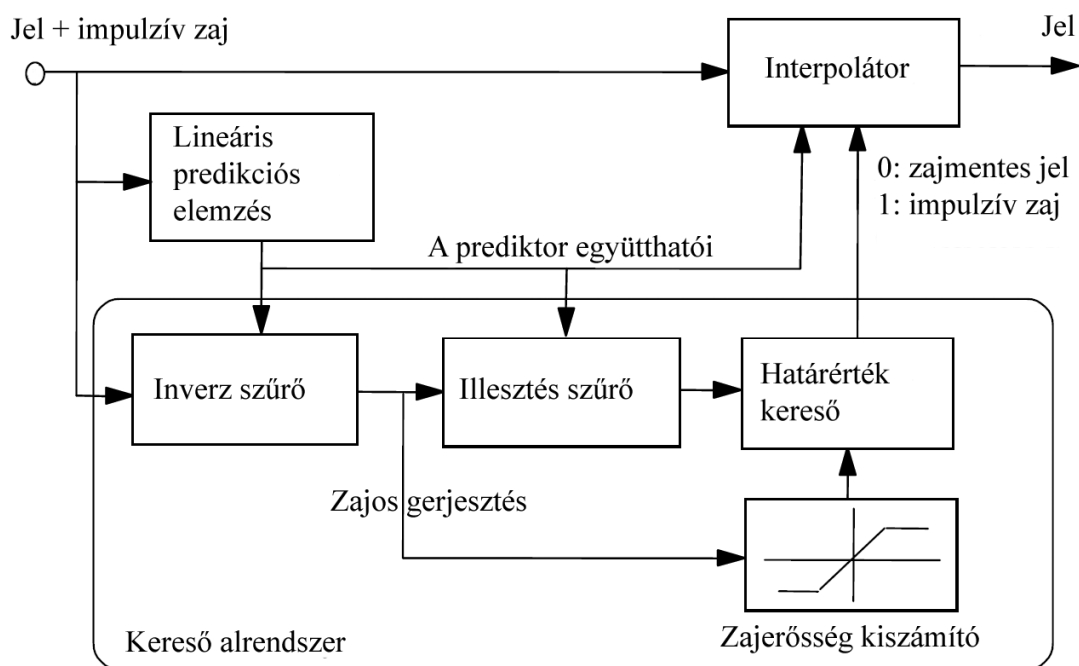
$\theta(m)$  egy alkalmazkodó határérték, mely erőteljes becsléssel határozza meg



$|y(m) - y_{\text{med}}(m)|$  átlagértékét,  $k$  pedig egy hangoló paraméter. A medián szűrők nem optimálisak, mert nem használhatók hatékonyan a jelgenerálás fiziológiájának előismerete vagy a jel és a zaj statisztikai eloszlásának modellezése nélkül.

### 3.3.4.3. Impulzív zaj eltávolítása lineáris prediktív modell segítségével

Az impulzív zavarok általában a jel egészét vizsgálva csupán csekély  $\alpha$  hányadát szennyezik. Mivel a jel jelentős  $1-\alpha$  része impulzív zajjal nem szennyezett, ezért előnyösebb megkeresni a zajos részeket (egyedülálló impulzusokat), és csak a torzult mintákat javítani. Ez a stratégia kiküszöböli a felesleges feldolgozásokat, és a minták egy viszonylag nagy része – mely nem szennyezett impulzív zajjal – érintetlenül eredeti állapotában marad. A következő ábra egy ilyen rendszer blokkdiagramját mutatja, mely két alrendszerből áll, egy keresőből és egy interpolátorból. A



29. ábra Egy impulzív zajeltávolító blokkdiagramja.

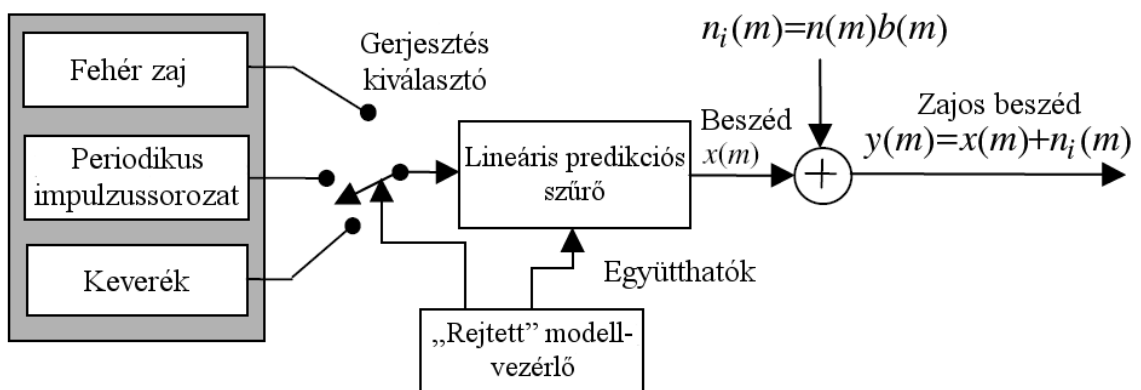
detektor megkeresi az impulzusok helyét, az interpolátor pedig kicseréli a torzult mintákat, figyelembe véve a torzult minta mindkét oldalán elhelyezkedő további mintákat. A detektor rész egy lineáris predikciós elemzőből, egy illeszkedésvizsgáló szűrőből és egy határérték keresőből áll. A kereső kimenete egy kétállású kapcsoló, mely az interpolátort vezérli. Impulzív zaj hiányában az interpolátor nincs bekapcsolva. Impulzív zaj jelenléte esetében az interpolátor bekapcsol és kicseréli a zaj által törölt mintákat.

## 3.3.4.3.1. Az impulzív zaj keresése

A legegyszerűbb módszer az ilyen jellegű zajok megtalálására, hogy meghatározzunk egy határértéket, és a határérték fölé eső mintákat zajnak osztályozzuk. Ez a módszer tökéletesen működik nagyon nagy impulzusok esetében, de ha a zaj amplitúdója a megadott határérték alá esik, akkor nem találja meg. A keresés eredményesebbé tehető, ha kihasználjuk a jel és az impulzív zaj jellemzőinek különbségét. Az impulzív zaj, vagy rövid időtartamú hullámvész bevisz egy a jelre nem jellemző megszakítást a jelbe. A megszakítás könnyebben megtalálható, ha a jelet szétválasztjuk. A szétválasztás (vagy digitális jelnél megkülönböztetés) művelete egyenértékű a kölcsönösségi viszony megszüntetésével, vagy a spektrális fehérítéssel. A lineáris prediktor modellezi a jel kölcsönösségi struktúráját, és ezen összefüggések megszüntetését végzi az inverz szűrő. A zajmentes  $x(m)$  jel lineáris predikciós modell segítségével a következő összefüggéssel írható fel:

$$x(m) = \sum_{k=1}^P a_k x(m-k) + e(m), \text{ ahol } a=[a_1, a_2, \dots, a_P]^T \text{ a lineáris prediktor } P \text{ rendű}$$

együttható vektora, és az  $e(m)$  gerjesztés egy zajhoz hasonló jel, vagy vegyes véletlenszerű zaj keveréke, és hullámvészok kvázi periodikus sorozata, melyet a következő ábra mutat. Az impulzív zajkereső nem a korrelálatlan kétállapotú impulzív zaj



30. ábra Zajos beszédmodell. A jelet egy lineáris prediktor modellezi, az impulzív zaj modellezésére pedig egy kétállapotú amplitúdómodulált modellt használunk.

alapján dolgozik, hanem azon a megfigyelésen alapszik, hogy a lineáris prediktorok jó modellt biztosítanak korrelált jelekre. A zajos  $y(m)$  jel a prediktor gerjesztett jelévé történő alakítása a következő hatásokat váltja ki:

- a.) Az jel amplitúdójának léptéke csökken az eredetileg gerjesztett jeléhez képest, mivel a zaj amplitúdójának léptéke változatlan marad, vagy növekszik.

- b.) A jel nem lesz kölcsönösen összefüggő, mert az impulzív zaj elmaszatólódott és átalakult léptékes alakba az inverz szűrő impulzus válaszaként.

Mindkét hatás javítja az impulzív zaj megtalálhatóságát. A beszéd és a zene hang időtartamának rezonanciája vagy a hangszer által erősített és spektrálisan alakított gerjesztésekből áll össze. Ez a gerjesztés véletlenszerűbb a beszédnél, és gyakran sokkal kisebb az amplitúdó-tartománya. Az impulzív zaj megtalálhatósága javítható az inverz szűréssel, mely lényeges lehet és függ a jel időben változó korrelációs struktúrájától. Ez a módszer egy optimális határértékképző használatával képes hatékonyan csökkenteni az impulzív zaj keresésének problémáját, a sorból kilógó minták megtalálását.

#### 3.3.4.3.2. Az impulzív zaj megtalálhatóság-javításának elemzése

A következőkben az impulzív zaj megtalálhatóságának javítását elemezzük ez inverz szűrő kimenetének vizsgálatával. Az előző egyenlet használatával a zajos jel modelljét a következő egyenlet segítségével (az előző átírt változata) írhatjuk fel:

$$y(m) = x(m) + n_i(m) = \sum_{k=1}^P a_k x(m-k) + e(m) + n_i(m), \text{ ahol } x(m), y(m) \text{ és } n_i(m) \text{ rendre a}$$

zajos jel, a jel és a zaj. Egy számítást használva  $\hat{\mathbf{a}}$  lesz  $\mathbf{a}$  prediktor együttható vektora, és a zajos  $y(m)$  jel inverz szűrhető és átalakítható  $v(m)$  zajosan gerjesztett

$$\text{jellé így: } v(m) = y(m) - \sum_{k=1}^P \hat{a}_k y(m-k) = x(m) - n_i(m) - \sum_{k=1}^P (a_k - \tilde{a}_k)[x(m-k) + n_i(m-k)],$$

ahol  $\tilde{a}_k$  a prediktor együttható számítási hibája. A zajos jel modelljének egyenlete ez alapján az egyenlet alapján következő egyenleté írható át egy másik alakba:

$$v(m) = e(m) + n_i(m) + \sum_{k=1}^P \tilde{a}_k x(m-k) - \sum_{k=1}^P \hat{a}_k n_i(m-k). \text{ Ebből az egyenletből három}$$

tétel határozható meg, mely elősegíti a zaj megtalálását a gerjesztési szekvenciában.

- a.) az  $n_i(m)$  impulzív zavar, mely általában domináns;
- b.) a régebbi  $P$  darab impulzív zajminta hatása az inverz szűrő működése végett, a jelenbe maszátolva:  $\sum \hat{a}_k n_i(m-k)$ ;
- c.) a gerjesztett jel varianciájának növekedése a paramétervektor számítási hibái miatt, mely a következő kifejezéssel írható le:  $\sum \tilde{a}_k x(m-k)$

Az inverz szűrő eredményeinek javítása a következő formulákkal írható le. Az impulzív zaj a zajos jelhez viszonyított aránya az alábbiakban ismertetett módon

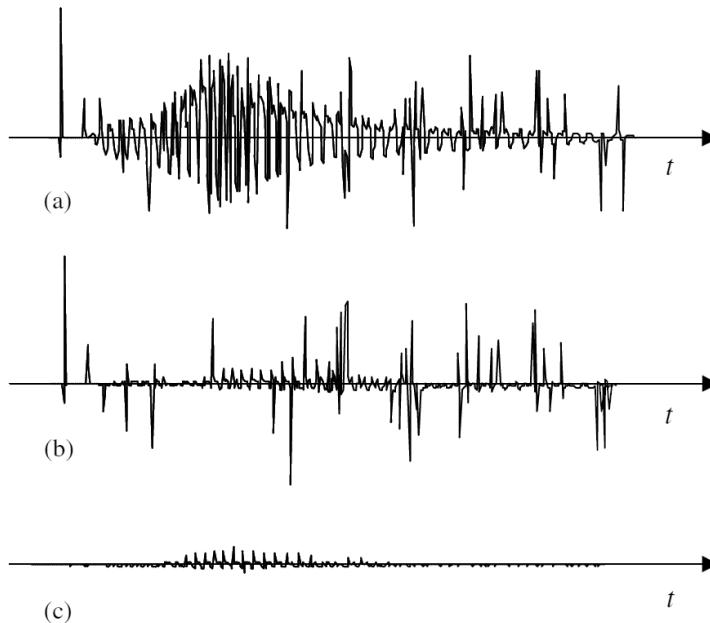
adható meg:  $\frac{\text{impulzív zaj energiája}}{\text{jel energiája}} = \frac{E[n_i^2(m)]}{E[x^2(m)]}$ , ahol  $E[\cdot]$  a várható érték operátora.

Az impulzív zaj detektálásában az érdeklődés tárgyát képező jel az impulzív zaj lesz, melyet a kísérő jelből kell megtalálni. Feltesszük, hogy a domináns zaj a  $v(m)$  zajos gerjesztési jelben az  $n_i(m)$  impulzív zaj, a két jel aránya a következőképpen írható fel:  $\frac{\text{impulzív zaj energiája}}{\text{gerjesztési energia}} = \frac{E[n_i^2(m)]}{E[e^2(m)]}$ . A teljes impulzív zaj jellel szembeni

aránya az előző két egyenlet hányadosából kapható meg, melyet a következő egyen-

lettel tudunk leírni:  $\frac{E[x^2(m)]}{E[e^2(m)]} = \text{nagyobbodás}$ .

Ez az egyszerű elemzés megmutatta, hogy az impulzív zaj megtalálhatósága függ a teljesítményerősítés karakterisztikájától, ennek következtében a lineáris predikciós modell rezonanciáitól. Beszéd esetében a jel amplitúdója zajmentes gerjesztett jelről lévén szó  $10^{-1}$  és  $10^{-4}$  között változik, ami a beszéd maga, ezért jelentős javulás érhető el az impulzív zaj megtalálhatóságában az inverz szűrő alkalmazásával a zajos beszédre. A következő ábra az impulzív zaj megtalálhatóságának



31. ábra Az inverz szűrő hatása a impulzív zaj megtalálhatóságára. (a) impulzív zajjal szennyezett beszéd, (b) az impulzív zajjal szennyezett beszéd gerjesztése, (c) beszéd gerjesztés impulzusmentesítve.

javítására alkalmazott inverz szűrő hatását mutatja. Az inverz szűrés hatása, hogy az  $x(m)$  jelet korrelálatlan  $e(m)$  jelgerjesztéssé alakítja, míg az impulzív zaj az inverz szűrő impulzív  $[1, -a_1, \dots, -a_p]$  zajátvitelének megfelelően torzul és megnövekszik, ahogy azt az iménti tétel mutatja a következő összefüggés alapján:

$\sum \hat{a}_k n_i(m-k)$ . Feltételezük, hogy a gerjesztés egy

normális eloszlást követő fehér zaj, egy szűrő, melynek együtthatói az inverz szűrő együtthatóival egyeznek képes növelni a torzult impulzív zaj megtalálhatóságát a gerjesztési jelben.

## 3.3.4.3.3. Kétoldali prediktor az impulzív zaj megtalálására

Az impulzív zaj megtalálásának egyik hatékony módja, hogy a jelet korrelálatlanná tesszük. Ez a módszer felhasználható a kétoldali lineáris predikciós modellben is. Egy  $x(m)$  minta kétoldali lineáris predikciója  $P$  darab régebbi (múltbeli) és  $P$  darab jövőbeli minta alapján dolgozik, és a következő egyenlettel definiálható:

$$x(m) = \sum_{k=1}^P a_k x(m-k) + \sum_{k=1}^P a_{k+P} x(m+k) + e(m),$$

ahol  $a_k$  értékek a kétoldali prediktor együtthatói, és  $e(m)$  a gerjesztési jel. Minden elemzés arra az esetre készül, mikor csak egy oldalról vizsgáljuk a jelet, de ez kiterjeszthető a kétoldali modellre is. Azonban a gerjesztés mértéke a kétoldali modell esetében kisebb varianciát mutat, mint az egyoldali modellben, mivel az imént leírt egyenletben mind a múltbeli és mind a jövőbeli összes mintára modellezett a minták korrelációja. Habár az iménti egyenletben egy nem véletlenszerű szűrő van, az inverze szükséges a kereső alrendszerbe, mely már véletlenszerű. A kétoldali prediktor használata további javulást eredményezhet az impulzív zaj megtalálhatóságában.

## 3.3.4.3.4. Az eldobott minták interpolációja

Az impulzív zaj hatására a minták visszavonhatatlanul torzulnak, és eltávolításukkal „űr” keletkezik, melyet interpolációval próbálhatunk kitölteni. A hiányosságok interpolálására – ahhoz, hogy ne maradjon hallható nyomuk –, magas hanghűségű interpolátor szükséges. A legkisebb négyzetes autoregresszív (LSAR) módszer magas minőséget eredményez viszonylag kis számú az impulzív zaj miatt elveszett minta esetében. Az LSAR interpolációs módszer egy kétszintű folyamat. Az első szinten az impulzív zaj mindkét oldalán megmaradt minták alapján becsléssel meghatározzuk a jel predikciós modelljének paramétereit. A második szinten a meghatározott modellparaméterek és a hézag két oldalán található minták alapján interpoláljuk az elveszett mintákat. Ezzel az interpolációs módszerrel az impulzív zaj által eltorzított audiojelek magas minőségben visszaállíthatók.

## 4. Zajsűrés a gyakorlatban

### 4.1. Zenekari stúdiómunka

#### 4.1.1. A felvétel

Egy hangstúdióban legyen szó akár könnyűzenei vagy komolyzenei műről a hangrögzítés során számos problémával kell szembenéznünk. Amennyiben megoldható, az akusztikus (nem elektromos) hangszerek felvételét úgynevezett süket-szobában célszerű kivitelezni, a visszhanghatások elkerülése végett, melyek a hanganyag torzulását okozhatják. A felvételt adott esetben csatornákra bontva külön-külön rögzítjük. Ez azt jelenti, hogy az egyes hangszerek külön játsszák a darabot, majd a felvételt követően egy keverőpult segítségével születik meg a végleges hanganyag.

A felvétel során is felütik a fejüket a zajok. Amivel mindenképpen számolni kell, hogy a bemikrofonozott jelforrás valamilyen távolságra helyezkedik el a mikrofontól, néha a zenész átéléssel játszik, és a hangszerrel együtt mozog. A távolság és az alkalmazott mikrofon minőségének függvényében a rögzítő technika mindenképpen rögzít egy kevés környezeti zajt is. Ezen zajok ellen a környezet megfelelő megválasztásával lehet a legkönnyebben védekezni.

További problémákat okoznak a mikrofonok viszonylag alacsony jelszintjei, és a jeleket továbbító kábelek. A kábeleket célszerű lehetőség szerint minél vastagabbra, tiszta anyagból készültre, és minél rövidebbre választani, a lehető legkevesebb csatlakozóponttal, ugyanis ez tudja a jel legjobb minőségű szállítását elvégezni. A nagy keresztmetszet és a viszonylag kis hossz biztosítja a kis ellenállást, mely a jelveszteséget csökkenti, és a tiszta anyag (például oxigénmentes réz, négy kilences tisztasággal – 99,99 % réz) pedig a szennyezések miatt kialakuló örvényáramok okozta zajtól óvnak bennünket. A csatlakozópontok további torzítást és zajt eredményezhetnek, az anyagok különbözősége, és a csatlakozó felületek mérete miatt. További eredményjavulás várható az arany, vagy felületén aranyozott kábelek alkalmazásától a magas frekvenciájú jelek átvitele szempontjából – a skin effektus miatt. A skin effektus azt jelenti, hogy a váltakozó áramot szállító vezetékben az elemi töltéshordozók a vezetőben annál távolabb haladnak a vezető középpontjától, minél nagyobb az átvitt jel frekvenciája.

Az alacsony jelszint miatt a digitális feldolgozhatóság elérése érdekében előerősítő fokozatokat kell a rögzítő készülék és a jelforrás(ok) közé építeni, melyek esetében az alkatrészek kialakítástól és minőségétől függően termikus zajgenerátorként működnek. A sávátviteli problémákat többnyire nem az erősítőelemek okozzák, hanem a mikrofonok.

További érdekesség a jelforrás szempontjából, hogy az elektromos hangszereket, mint például egy elektromos gitár, vagy basszusgitár általában egy úgynevezett alap erősítőre csatlakoztatjuk, majd ezen erősítő hangszórója elé helyezünk el mikrofont. Általában e hangszerek esetében az effekteket már az erősítő előtt, vagy annak effektútjába csatlakoztatjuk. Ilyenkor egy torzító effekt képes a hangszer hangszedője (pickupja) által felfogott zajokat nagymértékben felerősíteni, amit még – egy a torzító előtt elhelyezett (a játéktechnikát könnyítő) – dinamika kompresszorral lehet „növelni”. Ezek a zajok elektromágneses zajok, és sajnos a hangszer működéséből adódóan nehéz ellenük védekezni, mivel a húrok rezgését is ilyen módon érzékeli a hangszedő. Ilyen hangszerek esetén már két termikus zajgenerátor rendszer kapcsolódik a rögzítő elé, ezzel fokozva a zajszennyezést.

A rögzítő technika manapság már többnyire digitális rendszer, melynek az érkező jeleket egy ADC segítségével digitálissá kell alakítania. A jel ekkorra már zajjal erősen szennyezett lesz, a korábbi berendezések felépítésének és minőségének függvényében. A digitalizálás a mai korszerű hangkártyáknak nem jelentenek problémát, mégis stúdió célokra külön gyártanak – magasabb árfekvésű, jobb minőségű – kártyákat. A mintavételezési frekvencia megválasztása ekkor az egyik fő probléma. Általában a végeredmény CD lemezre kerül, ahol szabvány írja elő a 44100 Hz-es mintavételezési frekvenciát. Vannak hangkártyák, melyek ilyen mintavételezési frekvencia mellett nem produkálnak lineáris sávátvitelt, ellenben 48000 Hz-en igen. A felvett jelet visszahallgatva, van különbség 48000 Hz-en és a 44100 Hz-en mintavételezett jelek között az első javára, bár azt gondolhatnánk, hogy a hallható tartományon már ezek a frekvenciák kívül esnek. Másik érdekes kérdés a kvantálási hossz. A mai világban már a megfelelő dinamika elérése érdekében 16 bitnél nagyobb értéket választunk.

#### 4.1.2. Az utómunka

Az utómunkálatok első lépése a felvétel során a jelre került zajok eltávolítása. Ez esetben, ha sávonként rögzítettünk, akkor minden sávra külön célszerű a zajt eltávolítani, egészen pontosan annak hatását csökkenteni, ugyanis a zajok vélet-

lenszerű mivolta nem teszi lehetővé teljes eltávolításukat. Az ilyen felvételek esetén azért kell külön kezelni a zajokat, mert egy mikrofon és egy elektromágneses hangszedő több zajgenerátorral együttműködve más és más jellegű zajokat eredményeznek. Ha például egy dobfelszerelésre gondolunk, akkor észrevehető, hogy a megfelelő minőségű jel érdekében célszerű a hangszer egyes darabjait külön be-mikrofonozni, adott esetben más típusú mikrofonokkal, tehát ekkor már egy mikrofonrendszerrel beszélünk. Az egyes mikrofonok hangját külön kell bontani hangszín szempontjából, esetleg sztereó felvétel esetében még a panoráma sem fog egyezni minden esetben (például: lábcintányér jobboldal, pergődob közép, kísérő cintányér baloldal, természetesen a zenész kezességét figyelembe véve lehet fordítva is). A hangszínt minden esetben az egyes részeken külön kell beállítani, ami függ a mikrofon minőségétől, távolságától, helyzetétől és állásától. Ez a mikrofonrendszer mikrofononkénti hangszín beállításokkal másfajta zajt állít elő, mint például az imént említett elektromos gitár hangszedője egy torzítóval. Ezért is fontos a sávonkénti felvétel. Ezek a zajok csak ritka esetben tartalmaznak impulzív jellegű zajokat (vagy ha az ilyen zajok megjelennek, azok megszüntetéséről még a felvétel előtt célszerű gondoskodni), ezért a zaj csökkentésére a legalkalmasabb módszer a spektrális kivonás. Fontos, hogy a zajszűrő program (algoritmus) olyan legyen, hogy a hangképet ne, vagy csak nagyon kis mértékben módosítsa, a legjobb eredmény elérése érdekében.

Ezt követően az egyes hangszerek még további effektek hozzáadásával színesíthetők, valamint itt nyílik lehetőség az utolsó hangszínbeállításra, ugyanis a felvételkor még csak nyersanyag keletkezik. Itt adott esetben az egyes sávokat további analóg keverőpultba küldhetjük, de léteznek olyan stúdióban alkalmazható programok, melyek képesek a többsávós felvétel kezelésére. Ilyen program például a korábbi (a 2003-ban az Adobe cég által megvásárolt Syntrillium cég fejlesztésében megjelent) Cool Edit, és a manapság helyette alkalmazott Adobe Audition program, melyek kereskedelmi forgalomban kaphatók, de van ingyenes időkorlátos próbaváltozatuk is. Ezek a programok nem csak többsávós hanganyag effektezésére, hangszínbeállítására alkalmasak, hanem az azt megelőző zajcsökkentésekre is.

Természetesen a megfelelő hangzás elérése érdekében a hanganyagot a hangfelvétel elkészülte után hetekkel célszerű keverni, majd a végeredményt többféle berendezésen visszahallgatni, a maximális minőség elérése érdekében. A megfelelő hangzás elérése után még egy dinamika kompresszor segítségével lehet az anyagot



„összetömöríteni”. A mintavételezési frekvencia és a kvantálási hossz a későbbi hanghordozó függvényében változhat a rögzítési értékekhez képest, de ez sajnos a jel torzulását és néha még további zajok megjelenését is okozhatja.

## **4.2. Hagyományos hanglemez, hangkazetta archiválása CD-re**

Minden archiválási folyamat az adott hanghordozón tártolt hanganyag digitalizálásával kezdődik. Ezt akár otthonunkban is megtehetjük, feltéve, hogy rendelkezünk a forrás hanghordozó megfelelő minőségű lejátszására alkalmas berendezéssel. A hangminőséget jelentősen befolyásolhatja a jelforrás, továbbá az is, hogy milyen eszközzel végezzük a digitalizálást, és hogy milyen kábelek kötik össze az adott berendezéseket. Itt is lehet környezeti elektromágneses zavarokra számítani.

### *4.2.1. Hanglemez archiválása*

#### 4.2.1.1. A lejátszó rendszer

A gyártók általában többféle megoldást alkalmaznak a lemezejátszó mechanikájának és a mechanikai mozgásokat elektromos jellé alakító részegységek kivitelezésére.

A mechanika leglényegesebb elemei:

- Lemeztányér, amely a lemez forgatásáért, megtartásáért felelős, továbbá sok esetben a hangkar mozgatásával is foglalkozik, az alsó mechanikus szerkezeten keresztül. Többnyire nagy tömegű, az egyenletes forgás megvalósítása érdekében (nagyobb tehetetlenség, így nem képes a motor gyors fordulatszám-változásainak követésére). Fontos a jó csapágyazás és kenés a könnyű futás érdekében.
- Motor, amely a lemeztányért hajtja, nagyon fontos, hogy állandó fordulatszámú legyen nagyon kicsi ingadozással, mely valamilyen általában erőzárás közlőművel kapcsolódik a lemeztányérhoz (ez esetben fordulatszáma sokkal nagyobb, mint a tányéré).
- Hangkar (esetleg a mozgató mechanikával együtt), mely hangszedő fejjel együtt a lemeztűt egy shellnek nevezett tartóban rögzíti, és gondoskodik a megfelelő tűnyomásról és a barázdakövetésről.

A lemeztányér meghajtására háromféle megoldást szoktak alkalmazni, az egyik a dörzshajtás, a másik a szíjhajtás – gyakorlatilag ez az egyik legelterjedtebb –, és végül a direkt hajtás – ez közvetlen a motor forgórészére rögzített viszonylag

könnyű lemeztányért jelent. A motor lehet váltakozó áramú, mely többnyire a szabványos hálózatban alkalmazott frekvenciára épít a forgás tekintetében, vagyis szinkron motor, és lehet úgynevezett DC (egyenáramú) motor is. A közvetlen hajtású lemezjátszók esetében a legpontosabb fordulatszám elérése végett kvarcvezérlést alkalmaznak. Egyes típusoknál egy stroboszkóp lámpa segíti a fordulatszám beállítását, egy potenciométer beiktatásával, melyet általában az egyenáramú motorral szerelt típusok esetében alkalmaznak. A három megoldás közül a kvarcvezérlésű direkt hajtás a legpontosabb megoldás.

A hangkart mechanikusan áthelyezni képes lemezjátszók a lemeztányér alatt elhelyezett mechanikus szerkezet segítségével teszik mindezt. Létezik belőlük fél-automata és automata kivitel is. A hangkarban helyet foglaló (sok esetben leszerelhető, cserélhető) shell biztosítja a lemeztű számára kialakított hangszedő fej rögzítését. A szögkiegyenlítés végett – nem gyakran ugyan – alkalmaznak úgynevezett tangenciális hangkart, melynél nincs ilyen jellegű probléma. A nem tangenciális esetben a fej valamekkora szöget zár be a hangkarral, melyet vagy a hangkar meghajlításával, vagy a shell kialakításával érnek el.

Hangszedőből háromféle alaptípus létezik, melyek működésükben különböznek egymástól.

Az egyik – régebben viszonylag olcsó és hatékony megoldást biztosító – típus a kristály hangszedő. Ez működését tekintve a gyémánt vagy zafír tű által a lemeztől letapogatott rezgéseket egy vékony cső segítségével egy piezokristályra vezeti, mely a rezgést elektromos jellé alakítja. Ez a letapogató rendszer viszonylag nagy túnyomással működik jól, és ez sajnos nem kíméli a mikrobarázdás lemezünket. Hangzás tekintetében nem túl jó, de kielégítő eredmény elérhető vele. Nagy előnye, hogy nem igényel külön korrekciós előerősítő fokozatot, viszont impedancia-illesztés szempontjából nagyimpedanciás bemenetet igen.

A másik szinte a legelterjedtebb, már jó hangzást biztosító megoldás a mozgó mágneses hangszedő. Ez esetben (sztereó kialakítást feltételezve) két egymással 90°-os szöget bezáró tekerccs között mozgatja a vékony cső végére rögzített mágnest a lemeztől letapogatott mechanikai rezgés. Túnyomás tekintetében változatos megoldásokkal találkozhatunk, mert van itt  $\frac{3}{4}$  grammtól kezdődően kristály hangszedőt megszegyenítő 5 grammos nyomást igénylő kialakítás is. Itt már kell egy – a köznapi életben alkalmazott erősítők nagytöbbségében az alkalmazás idő intervaluma alatt fellelhető – korrekciós előerősítő fokozat, mely a mély hangokat kiemeli, a közép hangokkal nem tesz semmit, és a magas hangokat pedig csillapítja. Erre

azért van szükség, mert a lemez előállításakor így rögzítik a jelet – a barázda szélességének csökkentése miatt (több hanganyag fér el egy lemezen). Viszonylag egyszerűen cserélhető az elkopott tű, melyet még manapság is forgalmaznak bizonyos helyeken.

A harmadik megoldás a mozgó tekercses megoldás. Ez csak fejfel együtt cserélhető, nagyon drága, és hangzásvilágát tekintve ez a legjobb minőségű. Az előerősítő fokozat ide is kell, de az erősítésének nagyobbak kell lennie az előzónél, mert a jel feszültsége csupán kb. az egytizede a mozgó mágneses hangszedőének. Létezik olyan erősítő, amely mindkét jelet képes fogadni. A megoldások közül a legjobb hangminőség a mozgó tekercses hangszedővel érhető el, de ára, és a szükséges további eszközök miatt inkább a mozgó mágneses hangszedőt használjuk.

#### 4.2.1.2. Várható zajok

A hanglemezek mechanikus adathordozók, ezért nagyon sérülékenyek. A nem megfelelő tárolás következtében a por komoly károkat tud okozni a hangzásban. A nem megfelelő poreltávolítás apró karcokat okoz, melyek hallhatók lesznek lejátszáskor. A lemez felületére az ujjainkról került zsír is károsítja a hanghordozót, továbbá a felületen nagyobb karcok is kialakulhatnak (a nem megfelelően mozgatott hangkar következtében a tű is okozhat ilyen karcokat), melyek az impulzív zajokért felelősek.

A hanglemez kialakítása során a felületi egyenetlenségek keletkeznek, melyek szintén hallhatóvá válhatnak lejátszáskor. Ezért van különbség a keleti és a nyugati nyomású lemezek között, mivel más a gyártás alapanyaga és némileg a technológiája is.

A lemezjátszó mechanikája is képes zajokat előállítani. A lemeztányér a csapágyazás és kenés függvényében valamilyen mértékben rezeg. Ez egy precízen csapágyazott motor esetén egy direkt hajtású lemezjátszóban nem okoz problémát, mert általában a hallható tartományon kívül esik a rezgés frekvenciája.

Különbséget kell tenni a lemez száraz, illetve nedves lejátszása között is. A nedves lejátszás esetén a lemezt mosófolyadékkal lejátszás közben mossuk, így a szennyeződések eltávolítása is megtörténik, továbbá a folyadék vékony réteget képez a tű és a barázda információt tartalmazó részei között, ezért a hangzás is változhat nagyon kis mértékben.

A lemez kivitelezésénél fogva a korrekciós előerősítő fokozat is zajt termel, majd az azt követő illesztő fokozat viselkedik zajgenerátorként. Ezekben az erősí-

tőkben az aktív elemek melegszenek, termikus zajt előállítva. Továbbá a nagy erősítés és a mágneses működés miatt érzékenyek az elektromágneses zajokra is, mely megfelelő szigeteléssel orvosolható.

Tehát a lemezen megjelenik a helyreállítandó hézagokat maga után vonó impulzív zaj, és a spektrumát tekintve nemfehér zaj is, melyeket két különböző eljárással tudunk csökkenteni.

#### 4.2.1.3. Zajcsökkentési eljárás

A felvétel elkészítésekor figyelni kell az alkalmazott technikák jelszint és impedancia-illesztésére, valamint a mintavételezési frekvenciára és a kvantálási hossza is. Lehetőleg a cél adathordozóéval kompatibilis értékeket kell választani, a minőségromlás elkerülése érdekében.

A lemezjátszók készülhettek a keleti, illetve nyugati régióban is. A keleti régió általában az ötpólusú tuchel aljzatokat alkalmazta a jelek továbbítására, amely nem csak kialakításában, hanem jelszintjében és impedanciájában is eltér a nyugati régióban elterjedt (4 darab) RCA csatlakozókhöz képest. A tuchel csatlakozók kisebb jelszinten dolgoznak, és nagyobb bemeneti impedanciát igényelnek, továbbá kisebb felületen érintkeznek, tehát nem alkalmasak közvetlenül digitalizálásra. (Természetesen egy egyszerű erősítőfokozat közbeiktatásával alkalmassá tehetők a feladat elvégzésére az ilyen berendezések is.) A célszerű választás tehát a folyamat kivitelezésére egy *nyugati gyártású lemezjátszó, mozgó mágneses* hangszedővel, a megfelelő előerősítő fokozatot tartalmazó szintén *nyugati gyártású erősítővel*, mely illeszthető a szabványos hangkártyák vonalbemenetére.

A folyamat a digitalizálással kezdődik. A feladatra a számítógépbe célszerű egy eleve jó értékű jel–zaj aránnyal rendelkező, kis torzítású hangkártyát választani. A Creative cég régebben is készített ilyeneket, és ma is gyárt alkalmas eszközöket. (A Sound Blaster sorozat elemei.) A digitalizáláskor a jel erős zajszenyezéssel érkezik az eszközbe.

Az első lépés az impulzív zajok megkeresése és eltávolítása. Erre alkalmas az előző részben említett Adobe Audition program. Itt a határértékek feltérképezése után indulhat a folyamat, sajnos az eredmény csak a végrehajtás után lesz hallgatható, nincs „előzetes bemutatás” lehetőség. A folyamat az impulzív zaj mennyiségétől és a számítógép sebességétől függően akár több órát is igénybe vehet. Közben látható, hogy hány mintát vélt zajnak, és hány mintát sikerült kijavítania a programnak. Természetesen az javított/kihagyott arány a beállítás függvénye, na-

gyon karcos lemez esetén sok impulzív zaj van, melyek javítása eredményezheti a hangminőség romlását is, oly formában, hogy a magas frekvenciájú jeleket is impulzív zajnak tekinti program.

A következő lépés a spektrális kivonás. Ekkor mintát kell venni egy csak zajos részből, mely lehet a lemez bevezető része. A mintában szereplő pillanatképek száma beállítható, de nem feltétlenül eredményez jobb minőséget a nagyobb szám. Fontos paraméter a zaj csökkentésének mértéke, a precizitási szorzó és a simítás is. Ezek együttesen határozzák meg az eredmény minőségét. Van „előzetes bemutatás”, tehát meghallgathatjuk, hogy milyen eredményre számíthatunk, sőt meg-hagyható csak a zaj is.

A végén célszerű a jelet normalizálni, a hangerő (vagy jel–zaj arány) javítása érdekében. A mentés történhet szakaszonként is, ezzel a hanganyag darabolása is megtörténhet. Az eljárással akár 60–80 dB jel–zaj arány is elérhető, ami a kiindulási 30–40 dB-hez képest nem rossz eredmény.

#### 4.2.2. Hangkazetta archiválása

##### 4.2.2.1. A magnetofon és a hangszalag

A magnetofon vékony mágnesezhető réteggel bevont műanyag szalagon rögzíti a jeleket. A jelek, mivel az átmágnesezhető réteg vékony, nagyon kis feszültséget indukálnak a letapogató rendszerben, vagyis a magnetofon fejben, mely egy speciális kemény, kopásálló bevonattal rendelkező tekercs(pár). A kicsiny indukált feszültség erősítése egy korrekciós előerősítő fokozattal kezdődik, melynek a magas hangokat kell kiemelnie. A lejátszó rendszer ezt a szalagot húzza egyenletes sebességgel a fej előtt. A szalagsebességről egy kisebb átmérőjű fém és egy rugóval rá-szorított nagyobb átmérőjű gumigörgő gondoskodik, egyes esetekben ezt a rendszert megduplázzák, a stabilabb szalagsebesség érdekében. A görgők viszonylag könnyen szennyeződnek, ami a szalag nem kívánt oldalirányú mozgásához vezet; ez károsan befolyásolja a lejátszást. A szalagsebesség egy fontos jellemzője a berendezésnek, ami 4,76 cm/s egy átlagos magnetofon esetében. Ennek beállítása referenciaszalaggal és oszcilloszkóppal történik. A gyártók egy túrésen belülré állítják ezt az értéket, tehát két azonos szalagsebességű berendezést nehéz találni. A lejátszó és felvevő kombinált fej szennyeződhet, átmágneseződhet, ami szintén minőségromlást eredményezhet. Fontos paraméter még a szalagon található sáv és a fejen található érzékelő egymáshoz viszonyított elhelyezése is. Nem megfelelő beállítással a magasabb frekvenciájú hangok mennyiségének nagymértékű csökkenése

érhető el. A gyártók ezt is határértékek közé állítják, de létezik gyárilag rosszul beállított magnetofon is.

A szalag esetén problémák a szalagra felvitt mágnesezhető réteg minőségével kezdődnek, ebből alapjában véve négyféle típust alkalmaztak. Az első a normál vasoxid, a második a króm dioxid, a harmadik a ferrokróm (ezt már nem alkalmazták) a negyedik pedig a magnetit. A következő probléma a mágnesezhető réteg szemcsézete, egészen pontosan a mágneses információhordozók elrendezése. Ez a szalag gyártójától függően – azonos típusú mágnesezhető anyag esetén – anyagminőségben is eltérhet egymástól, tehát vannak jobb minőségű és rosszabb minőségű hangkazetták, annak függvényében, hogy a mágnességet mennyi ideig képesek megőrizni az egyes szalagok. A harmadik probléma a műanyag réteg vastagsága, mely lejátszáskor is okozhat problémát a szalagsebesség és az oldalirányú mozgás miatt, de tárolás esetén kérdéses, hogy az egyes átmágnesezett rétegek hogyan hatnak egymásra az idő múlásával. A műanyag hosszirányban képes a megnyúlásra, ami tovább fokozza a problémákat. Továbbá a vékony szalag nagyon érzékeny a mechanikai behatásokra, mint például a gyűrődés, amely durva minőségromláshoz vezethet. Az ilyen jellegű hiányok sem pótolhatók a jel folytonosságában, még digitális módszerrel sem. Nagyon fontos szerepe van még szalagok esetén az előmágnesezésnek, ugyanis ez is potenciális zajforrás.

#### 4.2.2.2. Várható zajok

A szalagos egységek esetében az előmágnesezési zaj után a szalagnak is van egy meghatározott értékű zaja, melyet egy olyan elektronika követ, ahol több termikus zajgenerátor is működik egy időben. A zaj ezért összetett, és általában egy elég jól hallható sistergés formájában jelentkezik, tehát főleg a magasabb frekvenciákon ölt jelentős mértéket.

A magnetofon esetén alkalmazható a Dolby zajcsökkentő rendszer, de lejátszáskor csak akkor, ha a felvétel is azzal készült, ellenkező esetben nem csak a zaj, hanem a magas hangok is eltűnnek. Ha nem tudjuk egy szalag esetén, hogy használtak-e valamilyen analóg zajcsökkentőt a felvételkor, akkor a berendezés zajcsökkentőjét kapcsoljuk ki, mert így a digitális zajszűrő hatékonyabban távolíthatja el a zajt, és több hasznos jel maradhat épen.

A magnetofonnal – a gyártók ígéretei szerint – akár 60 dB jel–zaj arány is elérhető a Dolby zajszűrővel, de sajnos a szalagok minősége és a közbeiktatott erősítőfokozatok miatt a digitalizálás után ez az érték 40–45 dB-re romlik.

Mivel vannak olyan rögzítő eszközök ebben a kategóriában, amelyek a törlést és egyben az előmágnesezést egy állandó mágnessel valósítják meg, ezért az előmágnesezés szintje közel sem lesz állandó, sokkal inkább jól hallható alacsony frekvenciájú előmágnesezési zaj lesz belőle. Ez a zajforma oly mértékben elnyújtott, és véletlenszerű, hogy algoritmikus csökkentése esetenként nem is lehetséges, ugyanis a zaj mintavételezése során nem választható ki úgy zajos rész, hogy e zajból kellő mennyiségű és minőségű jusson a mintába megfelelő kivonáshoz.

#### 4.2.2.3. Zajcsökkentési eljárás

A magnetofonszalagokról érkező jelek főként magas frekvenciájú zajt tartalmaznak, mely a korrekciós előerősítő fokozatból adódik. Ez a zaj mintavételezést követően spektrális kivonással csökkenthető a zavaró mérték alá.

Csak olyan esetekben ajánlott a magnetofonszalagok digitális hanghordozóra másolása, ha az anyag bír akkora fontossággal, hogy megtegyük (más hanghordozón, nem hozzáférhető). A magnetofonszalagokról tudjuk, hogy a normál szalag sávátvitele alig több mint a fele a krómdioxidos vagy magnetit szalagokéhoz képest, és a szalag zaja, ami a mágneses információhordozó minőségéből adódik, is lényegesen nagyobb az ilyen szalagoknál.

A magnetofonnál általában nem jelentkezik impulzív jellegű zaj, ezért a zajcsökkentés elvégezhető egyetlen lépésben. A digitalizált jelből (az elejéről vagy a végéről) zajmintát veszünk, majd pedig spektrális kivonással távolítjuk el a zajt. Az eredmény – bár az eljárás egy lépésből áll csupán – nem lesz olyan minőségű, mint a hanglemezek esetében, ugyanis a magasabb frekvenciájú hangokat a rögzítő rendszer és a szalag már korábban eltüntette, ezért lejátszáskor sem voltak már jelen. Ezek a hangok már semmilyen eszközzel nem állíthatók vissza.

## 5. Összegzés

Dolgozatom célja volt a környezetünkben előforduló zajok osztályozása, és azok csökkentésére megoldást találni. A potenciális zajforrások nagytöbbségére a dolgozatom első szakaszában világítottam rá. Tekintettel arra, hogy zaj hasonlóan az információt hordozó jelhez véletlenszerű, ezért modellezése és csökkentése közel sem olyan egyszerű, mint ahogyan azt a köznapi ember gondolja. A zajok modellezése a zajcsökkentés szempontjából jelentős, ugyanis egy megfelelő modell segítségével az eljárás könnyebben algoritmizálható, és alkalmazható a digitális jelfeldolgozásban. A számítógépek miniatürizálásával és gyors fejlődésével korunkban egyre nagyobb teret kap a kommunikációs és jelrögzítő rendszerekben a digitális jelek alkalmazása, mely a feldolgozhatóság és a hozzáadódó zajokkal szembeni jó „ellenálló-képességének” köszönheti sikerét.

A zajok modellezését kétféle szempont szerint közelítettem meg dolgozatomban, melyekből kiderült, hogy a véletlenszerű jelek előre meg nem jósolható értékeket is felvehetnek. A modellek alkalmazhatók más köznapi véletlenszerű események modellezésére is, melyek messze esnek az informatika tudományterületétől, mint például a közgazdaságtan, vagy a meteorológia, vagy az orvostudomány.

A probléma megoldása a zajcsökkentési eljárások alkalmazása. Többféle módszer létezik, melyekből csak néhányat volt alkalmam kiemelni, dolgozatom szűkös keretei miatt. A módszereket az adott problémának megfelelően, jelen esetben a zajok fajtáihoz illően kell alkalmazni, a várt eredmény elérése érdekében. A zajok csoportosítása, és forrásuk meghatározása ezért nagyon fontos, a legalkalmasabb eljárás kiválasztása szempontjából. Egy véletlenszerű jelsorozatból bizonyos esetben nehéz eldönteni még a szétválogatás után is, hogy melyik a számunkra hasznos információt hordozó jel. Erre nagyon szemléletes példa egy gépjármű motorja, mely az utas szempontjából zajforrás, de egy hozzáértő szerelő sok értékes információhoz juthat a hallott hangokból.

Az utolsó szakaszban az elméleti ismeretek gyakorlatba ültetését próbáltam kifejteni, saját tapasztalataim alapján. Több mint tíz évvel ezelőtt volt alkalmam több „házi stúdióban” készült könnyűzenei demófelvételhez segítséget nyújtani, és jómagam is aktív részese voltam egyes felvételek elkészítésének, mint hangszeres „zenész”. Jelenlegi iskolámban néha alkalmam nyílik kamatoztatni korábbi ismer-



reteimet, és dolgozatom megírásával nagyobb rálátást nyerhettem a későbbi munkámhoz a zajok forrásaira és megfelelő mértékű csökkentésükre.

A régebbi hanghordozókon tárolt, többnyire zenei felvételek korszerű adathordozóra mentése több ismerősömnek eszébe jutott már, de megfelelő eszköz hiányában maguk nem tudták elvégezni a műveletet. Érdeklődési körömbé tartozik a hangrendszerek felépítésének, és működésének megismerése és tanulmányozása, ezért több, mikrobarázdás lemez lejátszására alkalmas eszköz és magnetofon is található otthonunkban, mellyel a hanganyagok digitalizálásához forrást tudok biztosítani. A számítógépben található hangkártyákra mindig figyeltem, hogy ne csak gyenge minőségű hangjelet tudjon produkálni a komolyabb hangrendszer számára, hanem megfelelő működésével minőségi zenehallgatásra legyen alkalmas.

A dolgozat megírása során törekedtem arra, hogy a címben megragadott témát minél jobban, és több szempont szerint is kifejthessem, amit végeredményében sikerült megvalósítanom.

## 6. Irodalomjegyzék

1. Saaed V. Vaseghi – Advanced Digital Signal Processing and Noise Reduction  
Second Edition  
Copyright © 2000 John Wiley & Sons Ltd.
2. Valószínűségszámítás és statisztika előadások anyaga  
Nagy Márta – 2001–2002 tanév
3. Molnár-Sáska Gábor – Rejtett Markov modellek statisztikai vizsgálata  
Tézisfüzet – 2005  
forrás: [http://doktori.math.bme.hu/Molnar-Saska\\_tezisfuzet.pdf](http://doktori.math.bme.hu/Molnar-Saska_tezisfuzet.pdf)
4. dr. Tolnai János – Rádióamatőr vizsgára felkészítő tananyag  
Puskás Tivadar Távközlési Technikum, 2006.  
forrás: [http://www.puskas.hu/r\\_tanfolyam/r\\_tananyag.html](http://www.puskas.hu/r_tanfolyam/r_tananyag.html)

Debreceni Egyetem  
Informatika Kar

# Szakdolgozat

## Zajok és eltávolításuk

**Témavezető:** Dr. Papp Zoltán  
egyetemi adjunktus

**Készítette:** Gambár Péter István  
III. informatika tanár szakos,  
levelező tagozatos hallgató

Debrecen  
2009.

# Tartalomjegyzék

1. Bevezetés.....	4
1.1. A cím magyarázata.....	4
1.2. Témaválasztás indoklása .....	4
2. Zajok és modellezésük .....	5
2.1. Zajok fajtái és forrásai.....	5
2.1.1. Általános felosztás .....	5
2.1.2. Fehér zaj.....	7
2.1.3. Színezett zaj .....	7
2.1.4. Impulzív zaj.....	8
2.1.5. Tranziens zaj impulzus.....	10
2.1.6. Termikus zaj .....	11
2.1.7. Schottky-zaj.....	12
2.1.8. Elektromágneses zaj .....	13
2.1.9. Csatornatorzítások.....	14
2.1.10. Zajmodellezés .....	15
2.1.10.1. Hozzáadódó fehér normális eloszlású zajmodell (AWGN).....	16
2.1.10.2. Rejtett Markov modell.....	16
2.2. Zajmodellek .....	18
2.2.1. Valószínűségi modellek.....	18
2.2.1.1. Véletlenszerű jelek és sztochasztikus folyamatok.....	18
2.2.1.2. Valószínűségi modellek .....	19
2.2.1.3. Állandó és nem állandó véletlen folyamatok .....	23
2.2.1.4. Véletlen folyamatok várható értéke.....	24
2.2.2. A rejtett Markov modell.....	29
2.2.2.1. Nem stacionárius folyamatok statisztikai modellje .....	29
2.2.2.2. Rejtett Markov modellek.....	31
2.2.2.3. A rejtett Markov modell betanítása .....	34
3. Zajsűrítési módszerek.....	36
3.1. Zajzár .....	36
3.2. Analóg valós idejű zajsűrítők.....	36
3.2.1. Dolby zajsűrítő.....	36
3.2.2. Zajérvénytelenítés.....	38
3.3. Digitális jelek feldolgozása.....	39

---

3.3.1. Átalakítások.....	39
3.3.1.1. Analóg–digitális átalakítás.....	39
3.3.1.2. Digitális–analóg átalakítás.....	40
3.3.2. A Wiener szűrő.....	42
3.3.2.1. A legkisebb négyzetek becslése.....	42
3.3.3. Zajcsökkentés spektrális kivonással.....	44
3.3.3.1. A spektrális kivonás módszere.....	45
3.3.3.2. A spektrális kivonó megvalósítása.....	47
3.3.4. Az impulzív zaj eltávolítása.....	50
3.3.4.1. Impulzív zaj matematikai modellje.....	50
3.3.4.2. A medián szűrő.....	51
3.3.4.3. Impulzív zaj eltávolítása lineáris prediktív modell segítségével ...	53
4. Zajszűrés a gyakorlatban.....	58
4.1. Zenekari stúdiómunka.....	58
4.1.1. A felvétel.....	58
4.1.2. Az utómunka.....	59
4.2. Hagyományos hanglemez, hangkazetta archiválása CD-re.....	61
4.2.1. Hanglemez archiválása.....	61
4.2.1.1. A lejátszó rendszer.....	61
4.2.1.2. Várható zajok.....	63
4.2.1.3. Zajcsökkentési eljárás.....	64
4.2.2. Hangkazetta archiválása.....	65
4.2.2.1. A magnetofon és a hangszalag.....	65
4.2.2.2. Várható zajok.....	66
4.2.2.3. Zajcsökkentési eljárás.....	67
5. Összegzés.....	68
6. Irodalomjegyzék.....	70

## 1. Bevezetés

### *1.1. A cím magyarázata*

Általában hangfelvételek (függetlenül attól, hogy milyen eszközzel készülnek és milyen adathordozóra) és kommunikációs rendszerek esetében felütik fejüket az additív zajok. Ezen zajokat a hasznos jelhez szokták viszonyítani, és sajnos adott esetben a hasznos jel nagyságát is meghaladhatják, mely nagyban károsítja a jelek értelmezhetőségét. Több esetben azonban lehetőség nyílik az ilyen zajok mértékének, erősségének csökkentésére, mind hagyományos, mind digitális rendszerben. A zajok eltávolításánál nagyon fontos, hogy a hasznos jel minél nagyobb hányada maradjon ép, értelmezhető, és lehetőleg torzításmentes. Zajforrástól függően a zajok csökkentése hagyományos úgynevezett analóg rendszerekben is megvalósítható, akár megfelelő alkatrészek és eszközök alkalmazásával is. Jelen dolgozatban a tárolt és szállított hangjelek zajainak felismerésével, fajtaival és azok csökkentésével foglalkozom.

### *1.2. Témaválasztás indoklása*

Több éve egy zenei általános iskolában tanítok, ahol esetenként szükség van a kórus munkájának megörökítésére nem csak képi, de hang formájában is. A hangfelvételek megfelelő minőségű elkészítése és tárolása érdekében a rögzítő technika és az előtte található átalakító rendszerek által termelt zajok eltüntetése adott esetben nagyon fontos. Célom, hogy a dolgozat megírásával iskolámban minél hatékonyabb munkát tudjak végezni ebben a tekintetben is.

Személyes érdeklődési körömbé tartozik a zenei hangfelvételek digitális feldolgozása, illetve az erre alkalmas berendezések és szoftverek megismerése is, tehát többek között ez is motivált a szakdolgozat témájának megválasztásánál.

## 2. Zajok és modellezésük

### 2.1. Zajok fajtái és forrásai

#### 2.1.1. Általános felosztás

A zaj definiálható olyan nem kívánt jelnek, mely megakadályozza (zavarja) egy másik jellel (hasznos jel) történő kommunikációt vagy a másik jel mérését. A zaj maga is egy információhordozó, a zajforrás szempontjából. Például egy autó motorjának zaja információval szolgál a motor állapotáról. A zajok igen sokfélék és változatosak lehetnek, kezdve a hangfrekvenciás akusztikus zajoktól a rádiófrekvenciás elektromágneses zajokig.

A másik említésre méltó probléma a hasznos jel kiindulási alakjától való nem kívánatos eltérése, torzulása. Ebbe a csoportba tartoznak a különféle visszhangok, felharmonikusok, és az elveszett jelerészetek.

A zajok és torzulások a kommunikációs és mérési rendszerek fő ellenségei, ezért e káros mellékhatások eltüntetése érdemel kiemelt figyelmet. A zajszűrést és torzítás kiküszöbölését főként a mobil kommunikációban, a beszéd felismerésben, orvosi jelfeldolgozásban, radar és hanglokátor rendszerekben alkalmazzák a köznap életben.

A zajok életünk minden területén, minden környezetben jelen vannak, különféle formában. Az zaj forrásától függően a következő nagy csoportok különíthetők el egymástól:

- a.) *akusztikus zaj*: leggyakrabban mozgásból, vibrációból (rezgésből), ütközésből származik, ezáltal ez a zajforma – eltérő mértékben ugyan – jelen van a mindennapi környezetünkben. Ezt a zajt legtöbbször haladó gépjárművek, légkondicionálók, számítógépek hűtőventillátorai, közlekedés, emberi beszélgetések, szél, eső, stb. okozzák.
- b.) *elektromágneses zaj*: minden frekvencián jelen van, különösen a rádiófrekvenciákon. Minden elektromos berendezés, például a rádió- és TV adóállomások, és vevőkészülékek állítanak elő jelentős elektromágneses zajt.
- c.) *elektrosztatikus zaj*: feszültség jelenléte okozza, áram nélkül vagy folyó árammal. Ennek a zajnak az egyik legjelentősebb forrása a fénycsóvilágítás.

d.) *csatornatorzítás, visszhang, és csillapítás*: az átviteli csatorna nem ideális karakterisztikája következtében történik. A rádiófrekvenciás csatornák, melyeket a mobil kommunikációban használnak, – mikrohullámú frekvenciákon – különösen érzékenyek a terjedés karakterisztikájára és környezetére.

e.) *feldolgozási zaj*: az a zaj, amelyet a jel digitális–analóg átalakítása eredményez, például a zaj digitálisan kódolt beszédben vagy hasonló jelekben, vagy az elveszett csomagok a digitális adatkommunikációs rendszerekben.

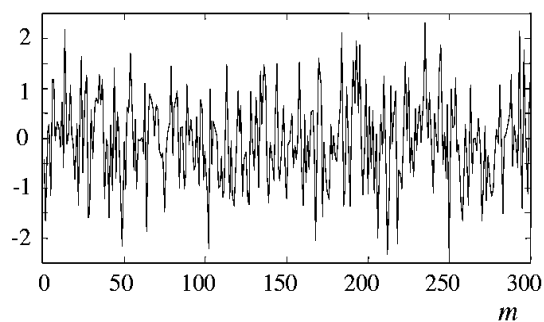
A zaj frekvenciájának vagy időkarakterisztikájának függvényében a következő kategóriákba csoportosítható:

- *Keskeny hullámsávú zaj*: olyan keskeny hullámsávon érkező zajhatás, mint az 50 Hz-es vagy 60 Hz-es brummogás (búgás) mely többnyire az elektromos tápegységekből, váltakozó áramú vezetékekből származik.
- *Fehér zaj*: tisztán véletlen zaj, melynek teljesen egyenes az energiaspektruma. A fehér zaj elméletileg az összes frekvenciát tartalmazza, azonos intenzitással.
- *Sávkorlátozott fehér zaj*: egyenes spektrumú zaj korlátozott sáv szélességgel, mely általában az alkalmazott eszköz korlátozott szélességű sávját, vagy az általunk használt sávot jelenti.
- *Szinezett zaj*: nemfehér zaj, vagy egyéb szélessávú zaj, melynek nem egyenes a spektruma, például: rózsaszín zaj, barna zaj, autoregresszív zaj.
- *Impulzusszerű zaj*: véletlenszerű rövid időtartamú impulzusokból áll, véletlenszerű amplitúdóval, hosszal és felbukkanással.
- *Tranziens zaj*: viszonylag hosszú időtartamú impulzusokból áll.



### 2.1.2. Fehér zaj

A fehér zaj definíció szerint egy összefüggéstelen jelekből álló zaj, mely minden frekvencián azonos energiájú. A zaj, melynek energiája minden frekvencián azonos  $\pm \infty$  tartományban, szükségképpen végtelen nagy energiájúnak kellene lennie, ennek következtében ez csak egy elméleti fogalom. Habár egy sávkorlátozott zaj,



1. ábra Fehér zaj

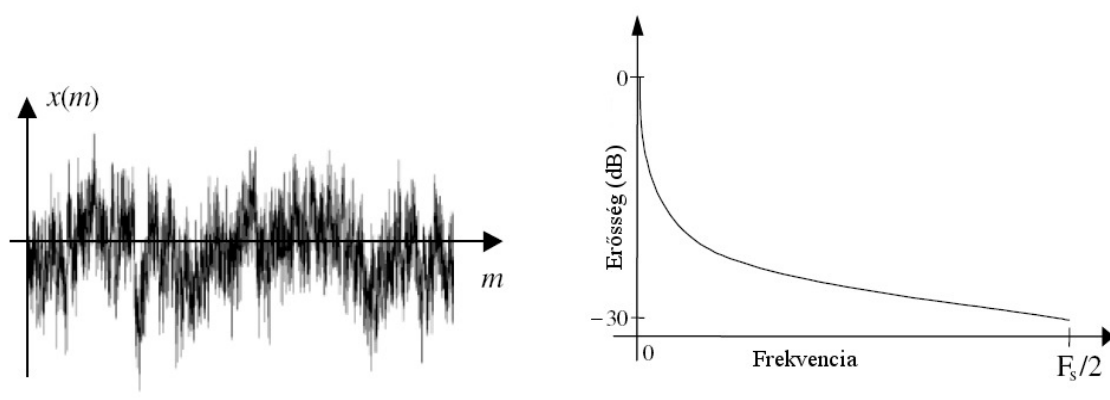
egyenes spektrummal – ha a zaj a teljes vizsgált sávot lefedi, amit a kommunikációs rendszer biztosít – a rendszer szempontjából minden tekintetben fehér zajnak számít. Például egy 10 kHz sávátvítellel rendelkező hangrendszer számára minden 10 kHz-nél szélesebb sávú egyenes spektrumú hallható zaj fehér zajnak

tűnik.

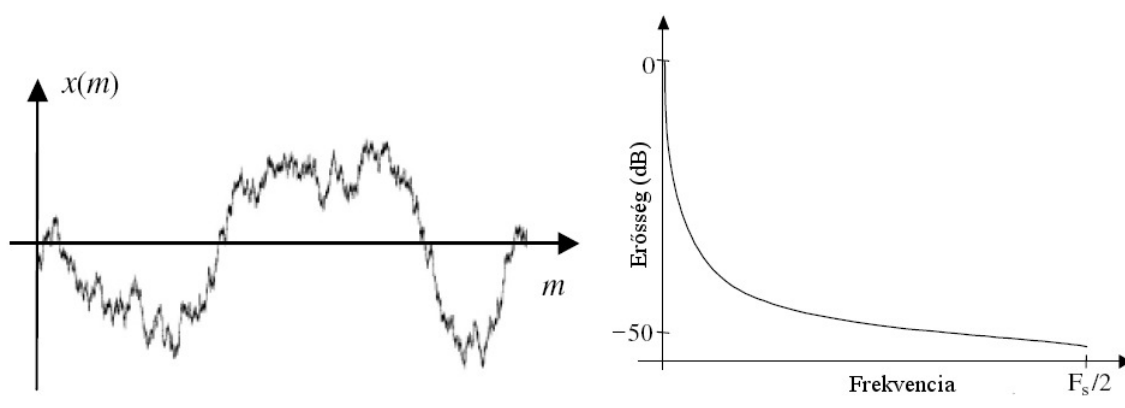
A tiszta fehér zaj egy elméleti fogalom, mivel végtelen energiájúnak kellene lennie a végtelen frekvenciatartomány lefedésére. Ráadásul a diszkrét időpontokban mintavételezett jelnek szükségképpen sávkorlátozottnak kell lennie, ahol a legnagyobb frekvenciaérték kisebb, mint a mintavételezési frekvencia fele. A sávkorlátozott fehér zaj egy sokkal gyakorlatibb fogalom, mely definíció szerint egy egyenes spektrumú zaj egy adott frekvenciasávban. Egy ilyen zaj energiája nem végtelen, ezért kiszámítható.

### 2.1.3. Színezett zaj

Bár a fehér zaj elmélete meglehetősen valós, és matematikailag kényelmes és használható közelítést ad több zajjal kapcsolatban, melyek felütik fejüket a telekommunikációs rendszerekben, sok egyéb zaj viszont nem fehér. A színezett zaj kifejezés bármely olyan szélessávú zajra utal, melynek spektruma nem a fehér zajéval egyező. Például a legtöbb hangfrekvenciás zaj, mint a mozgó autók zaja, a számítógépek ventilátorainak zaja, a fűrés zaja és emberi beszélgetés zaja, nemfehér, és túlnyomórészt alacsonyfrekvenciájú hangokat tartalmaz. Egy fehér zaj adott csatornán történő továbbításakor a zaj bizonyos frekvenciákon történő csillapodásával jellemezhető a „színezett zaj”, a csatorna spektrumának alakjától függően. A színezett zaj két klasszikus változata a rózsaszín (2. ábra) és a barna (3. ábra) zaj.



2. ábra Rózsaszín zaj és spektruma



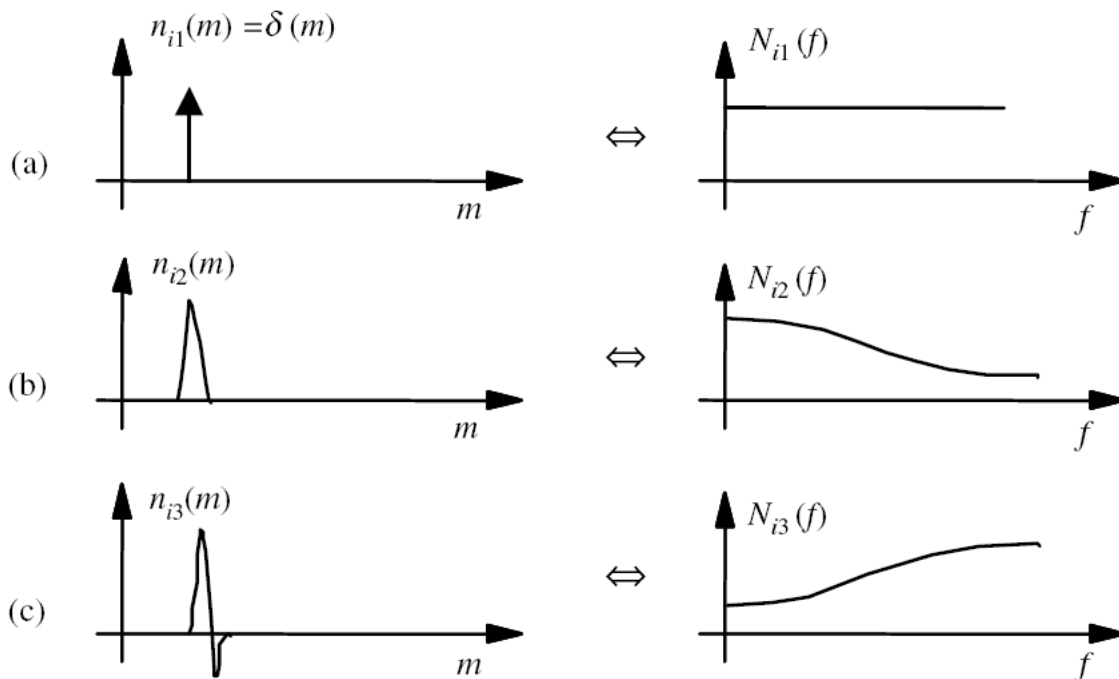
3. ábra Barna zaj és spektruma

#### 2.1.4. Impulzív zaj

Az impulzív zaj rövid időtartamú „be/ki” impulzusokból áll. Ezek a zajok többféle forrásból származhatnak, mint például kapcsolási zaj, kedvezőtlen csatorna-környezet, kimaradás vagy felületi kopás a hanghordozón, számítógép-billentyűzet kattogása, és így tovább. A 4. ábra (a) része egy ideális impulzív zajt mutat, annak spektrumával. A kommunikációs rendszerekben előforduló impulzív zaj jellemző paramétere a hossza, mely normális esetben több mint egy minta hosszúságú. Például hangrendszerek esetében a rövid időtartamú, éles impulzusok lehetnek akár 3 milliszekundum (60 minta 20 kHz-es mintavételezési frekvencia esetében) hosszúak, melyeket impulzív zajnak tekinthetünk. A 4. ábra (b) és (c) része két példát mutat rövid időtartamú impulzusokra és azok spektrumára.

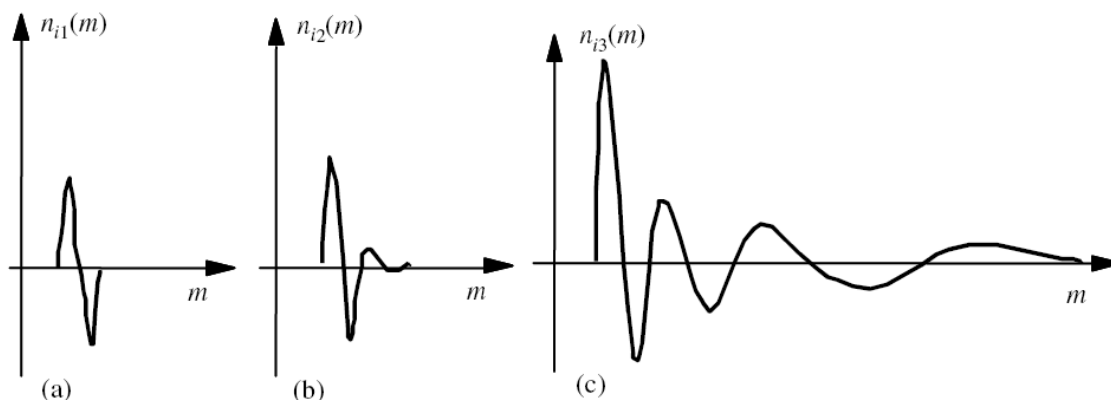
Egy kommunikációs rendszerben az impulzív zaj a térben és időben több pontból ered, azután a csatornán keresztül halad a vevő felé. A vett zaj időben szétszórt és a csatorna által alakított, és úgy kezelhető, mint a csatorna impulzus jelre adott reakciója. Általában a csatornák karakterisztikája lehet lineáris vagy nem lineáris,

időben állandó, vagy időben változó. Azonkívül sok kommunikációs rendszer válaszul a nagy amplitúdójú impulzusokra nem lineáris karakterisztikát mutat.



4. ábra Impulzív zaj

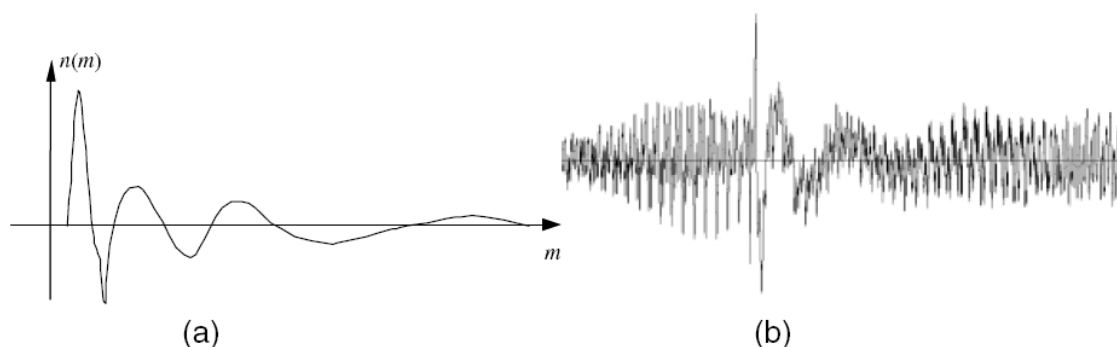
Az 5. ábra néhány példát mutat impulzív zajra, melyek tipikusan régi hanglemezeken fordulnak elő. Ebben az esetben a kommunikációs csatornának a lejátszó rendszert tekintjük, melyet időben állandónak feltételezhetünk. Az ábra három változatát mutatja a zaj-amplitúdóváltozásának az idő függvényében. Az 5. ábra (c) eleme mutatja, hogy egy nagy impulzslökés csillapodó tranziens impulzussorozatot hoz létre. Ezek a példák a lejátszó rendszer nem lineáris karakterisztikáját jelzik.



5. ábra Példák impulzív zajra

### 2.1.5. Tranziens zaj impulzus

Az tranziens zaj gyakran áll egy viszonylag rövid éles impulzusból, melyet egy csillapodó viszonylag alacsony frekvenciájú oszcilláció követ, ahogy azt a 6. ábrán is láthatjuk. A kezdő impulzust gyakran valamilyen belső vagy külső interferencia okozza, míg az oszcillációt gyakran a csatorna rezonanciája, melyet a kezdeti impulzus gerjeszt, és tekinthető a csatorna válaszának a kezdeti impulzusra. Egy telekommunikációs rendszerben az impulzív zaj a térben és időben több pontból származik, azután a csatornán keresztül halad a vevő felé. A vett zaj időben szétszórt és a csatorna által torzított, alakított és úgy tekintendő, mint a csatorna egy impulzusra adott válaszreakciója. Az impulzív zaj a kezdő impulzussal és az azt követő oszcillációval jellemezhető, a csatorna szempontjából.



6. ábra Egy karc impulzus és zene egy hanglemezről (b), hanglemez karc impulzusának átlagolt profilja (a)

A 6. ábra jól szemlélteti a tranziens zaj impulzus alakját, tekintsük meg a karcolás impulzust<sup>1</sup> az ábrán, mely egy sérült hanglemezről származik. A karcolás impulzusok a lemeztű és a hozzá kapcsolt elektromechanikus lejátszó rendszer akusztikus válasza a lemezen található fizikai megszakítások jelenlétére. Mivel a karcolás hatására előálló tranziens zajok tulajdonképpen a lejátszó rendszer impulzus válasza, elvárható, hogy az adott rendszer különböző helyeken található impulzusokra hasonló karakterisztikát mutasson. Ahogyan a 6. (a) ábra is mutatja egy tipikus karc impulzus hullámalak gyakran két különböző részre tagolódik:

- egy bevezető nagy amplitúdójú lökésre, mely a lejátszó rendszer válasza a rögzített médiumon (lemezen) található karcolásra, melyet
- csillapodó hullámzások követnek, további torzítást okozva ezzel.

<sup>1</sup> Karc(olás) impulzus alatt egy hanglemezen, a használat során kialakult (többnyire a sávval valamilyen szöveget bezáró) karc (tekinthető pillanatnyi jelfolytonosság-hiánynak is) hatására létrejövő lejátszáskor jelentkező ugrásszerű lökést érthetünk, mely közvetlenül a mechanikus jelérzékelőre hat, és a mechanikai–elektronikai átalakítást követően végighalad a teljes elektronikus rendszeren is.

A kezdő impulzus általában viszonylag rövid, és a hossza a 1–5 ms tartományba esik, míg a hullámzó farok hosszabb időtartama elérheti akár az 50 ms-ot vagy többet is.

A 6. (b) ábrán látható csillapodó hullám (farok) frekvenciája is csökken az idő múlásával. Ez a viselkedés az elektromechanikus lejátszó-rendszer nem lineáris válaszmechanizmusa a lemezen található karcokra. Tapasztalat alapján több hanglemezről származó recsenés zaja jól meghatározható hullámalakot mutat, és viszonylag kevés tipikus sablonnal jellemezhető.

### 2.1.6. Termikus zaj

A termikus zaj – más néven Johnson-zaj (felfedezőjéről J. B. Johnsonról<sup>1</sup> elnevezve) – hővel stimulált részecskék rendezetlen mozgásából adódik. A termikus zaj elméletének alapjai a termodinamikában keresendők. A termodinamikában a szabad részecskéket hőmérsékletfüggő rendezetlen mozgás jellemzi, mint például a szabad gázmolekulák egy tartályban, vagy a szabad elektronok egy vezetőben. Bár ezen rendezetlen mozgások átlagértéke közel nulla, az átlag körüli hullámzások alkotják a termikus zajt. Például a gázmolekulák rendezetlen mozgásai és ütközései egy behatárolt térben rendezetlen hullámzást okoznak az átlagos nyomás körül. A hőmérséklet növekedésével a molekulák vagy más elemi részecskék kinetikai energiája növekszik, és ezzel a termikus zaj is.

Ehhez hasonlóan egy elektromos vezetőben nagy számú szabad elektron van jelen (az ionokkal együtt), melyek rendezetlenül rezegnek az egyensúlyi pozíciójuk körül, és útját állják a további elektronok mozgásának. Az elektronok rendezetlen szabad mozgásai spontán áramokat, vagy más szóval termikus zajt képeznek, melynek átlaga nulla, mivel feszültség hiányában minden elektron mozgása különböző irányú (nincs egyirányú áramlás). A környezet hőmérsékletének növekedésével a vezető hőmérséklete is növekszik, mely az elektronokat magasabb energiaszintre viszi, ezzel növelve a rendezetlen áramfolyásokat. Egy fémréteg ellenállásban a pillanatnyi feszültségek négyzetének átlagértéke a termikus zaj miatt a következő összefüggéssel adható meg:  $\overline{v^2} = 4kTRB$ , ahol  $k = 1,38 \cdot 10^{-23} \frac{\text{J}}{\text{°K}}$  a Boltzmann-állandó,  $T$  az abszolút hőmérséklet °K-ben,  $R$  az ellenállás ohmban, és  $B$  a sáv szélesség. Az iménti egyenletből és az azt megelőző érvelésből egy áramkörü

<sup>1</sup> John Bertrand “Bert” Johnson (1887–1970): svéd születésű amerikai villamosmérnök és fizikus. Ő fejtette ki először részletesen a kábeleken közvetített információ véletlen interferenciájának alapjait. (Wikipédia – [http://en.wikipedia.org/wiki/John\\_B.\\_Johnson](http://en.wikipedia.org/wiki/John_B._Johnson))

lapba ültetett fémréteg ellenállás úgy viselkedik, mint egy termikus zajgenerátor, mely feszültségnégyzetének átlaga  $\overline{v^2}$ , és a belső ellenállása  $R$ . Az ily módon előálló termikus zajgenerátor teljesítménye számítható ki a következő összefüggéssel:

$$P_N = \overline{i^2} R = \left( \frac{v_{rms}}{2R} \right)^2 R = \frac{\overline{v^2}}{4R} = kTB [\text{W}], \text{ ahol } v_{rms}^1 \text{ az effektív feszültség értéke. A ter-}$$

mikus zaj spektrumsűrűsége megadható a  $P_N(f) = \frac{kT}{2} \left[ \frac{\text{W}}{\text{Hz}} \right]$  összefüggéssel. Ebből

az egyenletből kiderül, hogy a termikus zaj spektrumsűrűsége egyenes, azaz a termikus zaj egy fehér zajnak tekinthető (mely általában sávkorlátozott). Ezt az összefüggést különösen a magas rádiófrekvenciák ( $10^{13}$  Hz) esetében fontos.

### 2.1.7. Schottky-zaj

A Schottky-zaj<sup>2</sup> elmélete az elektroncsövek katódjának elektronkibocsátás-ingadozásából származik. Az különállónak tekintett elektronok áramlás során véletlenszerű időpillanatokban érkeznek, ezért az átlagérték körüli ingadozást okoznak a részecskefolyamban. Ezek az ingadozások a részecskefolyam sebességében Schottky-zajt okoznak. Egyéb példák Schottky-zajra a fotonok áramlása egy lézer fénysugárban, elektronok és lyukak áramlása valamint rekombinációja<sup>3</sup> félvezetőben, vagy a fotóelektronok árama, melyek a fotódiódákban emittálódnak. A véletlenszerűség jelenléte a részecske-kibocsátás vagy -érkezés mértékében magába foglalja, hogy a Schottky-zaj Poisson eloszlással modellezhető. A hosszú megfigyelési idejű kísérletek esetén az érkezések átlagérték körüli ingadozása a normális eloszláshoz közelít. Míg a termikus zaj a részecskék „nyugalmi állapotban” történő véletlenszerű mozgásából adódik, addig a Schottky-zaj egy (feszültséggel) irányított részecske-áramlásra értelmezhető.

<sup>1</sup>  $v_{rms}$  – root mean squared voltage: a feszültség effektív értéke

<sup>2</sup> Walter Hermann Schottky német fizikus, a film-rácsú vákuumcső (1915) és a tetróda (1919) feltalálója – munkáit a Siemensnél végezte. 1938-ban megfogalmazott egy elméletet, mely a Schottky-hatást jelezte, ma pedig a Schottky-diódákban használjuk. – [http://en.wikipedia.org/wiki/Walter\\_H.\\_Schottky](http://en.wikipedia.org/wiki/Walter_H._Schottky)

<sup>3</sup> „A félvezető nem csak szabad elektronjai útján vezeti az áramot (mint a fémek), hanem egy másik vezetési mechanizmus: a lyukvezetés útján is. Az eredetileg elektromosan semleges atom a negatív töltésű elektron távozása folytán pozitív töltésűvé válik, ezért a lyukat pozitív töltésűnek tekintjük. Előfordulhat, hogy egy szomszédos atom valamelyik vegyértékelektronja szintén elszabadul, és „beugrik” az előbb említett (pozitív töltésénél fogva az elektront vonzó) lyukba. Ezzel az előbbi lyuk (és az oda beugró szabad elektron is) megszűnik, hiszen a rácsszerkezet rendje az adott helyen helyreállt (ezt a folyamatot nevezik *rekombinációnak*), viszont a most elszabadult szabad elektron helyén új lyuk keletkezett. Ha valamelyik szomszédos atomból elszabaduló elektron most ebbe a lyukba ugrik be, tehát rekombinálódik, a lyuk a most elszabadult atom helyére »vándorol át.«” Forrás: Internet – [http://www.puskas.hu/r\\_tanfolyam/felvezetok.pdf](http://www.puskas.hu/r_tanfolyam/felvezetok.pdf), 2. oldal

Tekintsünk egy elektromos áramot, mint különálló elektromos töltéshordozók áramlását. Ha a töltéshordozók egymástól függetlenül mozognak, akkor az áram ingadozása megadható a  $I_{Noise}(rms) = \sqrt{2eI_{dc}B}$  összefüggéssel, ahol  $e = 1,6 \cdot 10^{-19}$  coulomb (egy elektron töltése), és  $B$  pedig a mérhető sávszélesség. Például egy  $I_{dc}$  1 amper „állandó” áram, és 1 MHz sávszélesség mellett, az effektív hullámzás 0,57 mikroamper lesz. Az iménti összefüggés feltételezi, hogy egymástól független töltéshordozók alkotják az áramot. Abban az eset mikor a töltéshordozóknak egy akadályon kell áthaladniuk, mint például egy félvezető dióda p–n átmenete, a töltéshordozók nem együtt mozognak; viszont nem igaz ez a fém vezetőkre, ahol a töltéshordozókat nagyfokú kölcsönhatás jellemzi.

### 2.1.8. Elektromágneses zaj

Virtuálisan minden elektromos eszköz, mely energiát állít elő, energiát fogyaszt vagy energiát szállít, elektromágneses zajok potenciális forrása. Az ilyen elektromágneses zaj interferenciának tekinthető más rendszerek szempontjából. Általában minél nagyobb a feszültség vagy az áram és minél közelebb helyezkedik el az áramkörhöz vagy elektromos eszközhöz, annál nagyobb lesz az áramkörben tapasztalható elektromágneses zaj mértéke. Az elektromágneses zajok általános forrásai a transzformátorok, rádió és televízió adóállomások, mobiltelefonok, mikrohullámú adóállomások, váltakozó áramú vezetékek, motorok és motorindítók, generátorok, relék, oszcillátorok, fénycsövek, és villámlások (zivatarok).

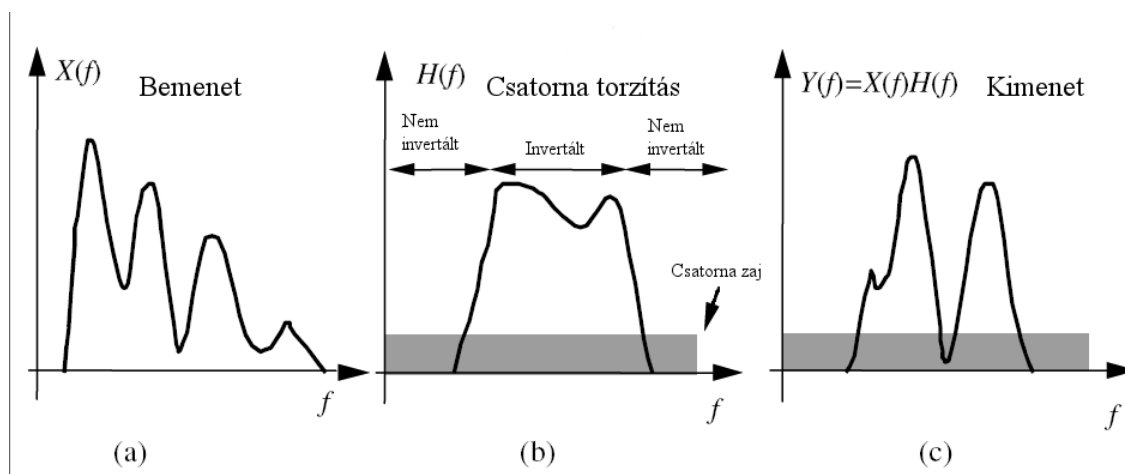
Az elektromágneses zajok két alapvető csoportba sorolhatók: elektrosztatikus zajok, és mágneses zajok. Ez két típus alapvetően különbözik egymástól, ezért más-más zajvédelem szükséges hozzájuk. Sajnos az imént felsorolt zajforrások legtöbbször a két típus kombinációját állítja elő, mely tovább bonyolítja a zajcsökkentési problémát.

Az elektrosztatikus tereket a feszültség jelenléte generálja, árammal, vagy anélkül. Az elektrosztatikus zajok egyik legáltalánosabb forrásai a fénycsövek. Mágneses terek keletkeznek elektromos áram hatására, vagy állandó mágnes jelenléte esetén. A motorok és a transzformátorok szolgálhatnak például az iménti felsorolásból, és hogy ne menjünk messzire a Föld mágneses tere is itt van környezetünkben. Ahhoz, hogy zajfeszültség indukálódjon egy vezetőben, a mágneses erővonalakat metszenie kell. A generátorok működési elve is ezen a jelenségen alapul. A váltakozó mágneses tér jelenléte – amilyen például az 50/60 Hz-es elektromos vezetékeket körülveszi – feszültséget indukál minden álló vezetőben, ahogy a mág-

neses tér kialakul, majd eltűnik (váltakozó mágneses tér). Ehhez hasonlóan, ha egy vezető mozog a Föld mágneses terében, zajfeszültség indukálódik benne, ahogy mozgás közben metszi a Föld mágneses erővonalait.

### 2.1.9. Csatornatorzítások

Ha egy jelet egy csatornán keresztül továbbítunk, a jel alakja változik, torzul a csatorna frekvenciaátvivő képességének és csillapítási paramétereinek megfelelően. A csatornatorzítás két fő megnyilvánulási formája a folytonossági torzítás és a fázistorzítás. A rádiókommunikációban létezik egy multi-út hatás is, mely szerint ha a jel több átjátszón megy keresztül, bizonyos esetekben különböző utakon haladva, ezzel a jel többféle változata érkezik a vevőbe különböző késleltetéssel és csillapítással. A csatornatorzítások csökkenthetők, vagy egészen egyszerűen szétválaszthatók a kommunikációs folyamatokban, ezért a csatorna modellezése és ki-



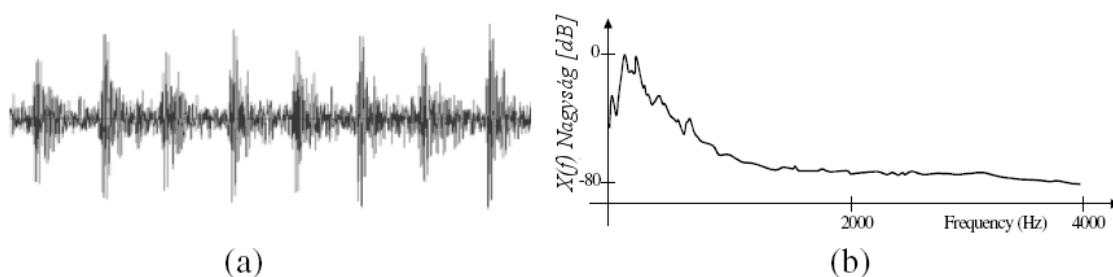
7. ábra Csatornatorzítás

egyenlítése nagyon jelentős összetevői egy modern digitális kommunikációs rendszer kialakításának. A csatorna kiegyenlítése különösen fontos a modern sejtes felépítésű kommunikációs rendszerekben mivel a csatorna karakterisztikák és terjedési paramétereik különbsége igen változatos, sokkal változatosabb, mint a vezetékes rendszerekben. A 7. ábra egy csatorna sávátviteli jellemzőit mutatja, mely két nem megfordítható és egy megfordítható régióra oszlik. A megfordítható csatornarégióban a jelek erősen torzultak elvesztek a csatornazaj miatt. A nem megfordítható csatornarégióban a jelek torzultak ugyan, visszanyerhetők. Ez a példa remekül mutatja, hogy a csatornafordító szűrőket nagy gondossággal kell elkészíteni, a nem kívánt eredmények elkerülése érdekében, mint például a zajerősítés az alacsony jel-zaj arányú frekvenciákon.



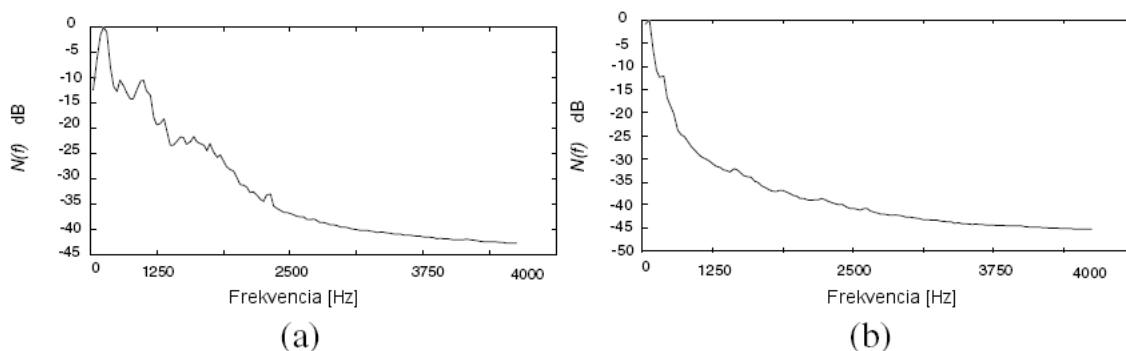
## 2.1.10. Zajmodellezés

A cél a modellezés esetén, hogy a jelben vagy zajban jelalakokat és mintákat jellemezzünk. A zaj pontos modellezéséhez egy olyan struktúrára van szükségünk, mely mind időben mind spektrumában jellemzi a zajt. A pontos zajstatisztika modellezéshez a kulcs a zajos jel megfelelő osztályozása és a hasznos jel erősítése. Ezen látszólag egyszerű jel/zaj osztályozási munka döntően függ a rendelkezésre álló hasznos jel mennyiségétől, erősségétől és a zajmodellektől, és ezen modellek alkalmazásától.



9. ábra A fúrózaj (a) idő-hullámformája (b) frekvenciaspektruma

A leghasználhatóbb szinte nélkülözhetetlen módszer az összes közül a zaj felépítésében való éleslátásra a Fourier transzformáció a zaj frekvenciaanalizálására. A 8. ábra egy elektromos fúró zaját mutatja, és ahogy az várható a zaj periodikus. Az ábra (b) része mutatja a fúró zajának spektrumát, melyen látható, hogy a zaj energiájának jelentős része a spektrum alacsonyabb frekvenciáira koncentrálódik. Tényszerűen igaz, hogy a legtöbb hangfrekvenciás jelnek és zajnak van túlnyomórészt alacsonyfrekvenciás a spektruma. Bár meg kell jegyezni, hogy a viszonylag kisebb energiájú magasabb frekvenciájú hangjelek fontos szerepet játszanak az élmény és a minőség átadásában. A 9. ábra (a) és (b) része példával szolgál az autók belsejéből származó zaj spektrumából (egy BMW és egy Volvo zajai láthatók). A



8. ábra Autók zajának spektuma (a) BMW 112 km/h, (b) Volvo 112 km/h

zajok nem statikusak, változatosak, és a következő forrásokból származhatnak:

- a.) kvázi periodikus zaj a motorból, és a jármű forgó alkatrészeiből,
- b.) zaj a kerekek és a talaj érintkezéséből és az út felületéből adódóan,
- c.) az autó körüli légáramlatok a légjáratokon, ablakokon, napfénytetőn keresztülhaladva,
- d.) az éppen előzőtt jármű zaja

Az autó zajainak jellemzői változnak a sebességgel, az út felületének állapotváltozásával, az időjárással és az autó környezetével.

A zajmodellezés legegyszerűbb módja, melyet gyakran használunk a jelenlegi gyakorlatban, hogy kiszámítjuk a zaj statisztikáját jelmentes időszakok alapján. Optimális esetben a Bayes féle jelfeldolgozási módszerekben egy készlet valószínűségi modell van előkészítve a jel- és zajfeldolgozásra. Ezek a modellek alapjául szolgálnak a jelek és zajok dekódolásának, zajos jelek felismerésének, erősítésének.

#### 2.1.10.1. Hozzáadódó fehér normális eloszlású zajmodell (AWGN<sup>1</sup>)

A kommunikációelméletben a gyakran feltételezzük, hogy a zaj hozzáadódó fehér normális eloszlást követő. Bár számos probléma esetében ez valós feltevés, és matematikailag kényelmes és használható megoldáshoz vezet, a gyakorlatban a zaj gyakran időben változó, kölcsönösen összefüggő, és nem normális eloszlást követő. Ez részben igaz az impulzív típusú zajokra, melyek nem állandóak, nem normális eloszlásúak, ezért nem modellezhetők az AWGN feltevés használatával. A nem állandó és nem normális eloszlású zajok modellezésére a Markov-féle sorozat alkalmas, stacionárius alfolyamatok leírásával.

#### 2.1.10.2. Rejtett Markov modell

A legtöbb zaj nem állandó, ez érvényes a zaj statisztikai paramétereire, úgy mint a zaj középértéke, szórásnégyzete, erőssége és spektruma, és ezek időbeli változása. A nem stacionárius zajok rejtett Markov modellek (HMM<sup>2</sup>) segítségével modellezhetők. Egy HMM lényegében stacionárius alfolyamatok véges sok állapotú Markov sorozata. Az állandó zaj modellezhető egy egyállapotú HMM segítségével. A nem állandó zaj modellezésére a többállapotú HMM alkalmas, amely képes a zaj időbeli változásának követésére, véges sok állandó állapottal. Nem normális eloszlású zajok esetében egy Gauss sűrűségi modell keverékét alkalmazhatjuk, a zaj terének modellezésére minden állapotban. Általában az állapotok száma modellen-

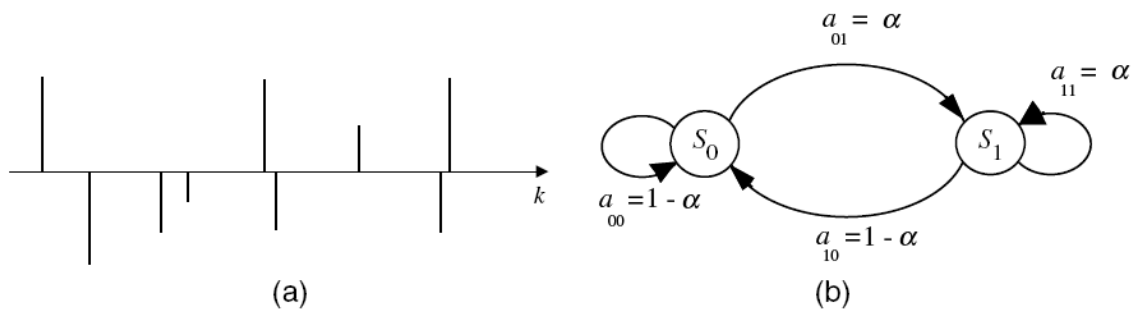
---

<sup>1</sup> AWGN: Additive White Gaussian Noise Model

<sup>2</sup> HMM: Hidden Markov Model

ként, és a keverékek száma állapotonként szükséges a zaj pontos modellezéséhez, mely a nem állandó zaj karakterisztikájától függ.

Egy példa a nem állandó zajra az impulzív zaj, melyet a 10. ábra (a) része szemléltet. Az ábra (b) részén az impulzív zaj szekvencia egy kétállapotú HMM modellje látható, ahol az  $S_0$  állapot jelképezi az „impulzus-ki” periódusokat az impulzusok között, és az  $S_1$  állapot jelképezi magát az impulzust. Azokban az esetekben ahol minden impulzus jól definiált átmeneti felépítéssel rendelkezik, előnyös lehet a többállapotú HMM magának az impulzusnak a modellezésére.



10. ábra (a) Impulzív zajszekvencia, (b) az impulzív zaj kétállapotú modellje

## 2.2. Zajmodellek

### 2.2.1. Valószínűségi modellek

A valószínűségi modellek alapja az információelmélet. Az információ maga is mennyiségileg a valószínűség logaritmusaként fejezhető ki. A valószínűségi modellek általában azzal foglalkoznak, hogy történések előfordulását megjósolják és jellemzésük véletlenszerű események alapján, különféle területeken, mint például telefonhívások számának megjósolása egy fővonalon a nap egy meghatározott időszakában, közúti forgalommodellezés, időjárás előrejelzés, pénzügyi adatmodellezés, előre jelezni a drogok hatását megadott adatok alapján az orvosi gyakorlatban. A jelfeldolgozásban a valószínűségi modelleket általában véletlenszerű jelek ingadozásának leírására alkalmazhatjuk, mint például mintafelismerés, jelkódolás és jelkiértékelés.

#### 2.2.1.1. Véletlenszerű jelek és sztochasztikus folyamatok

A jelek egyik legalapvetőbb jellemzőjük szempontjából két nagy csoportba osztályozhatók: determinisztikus (pontosan meghatározott) jelek és véletlenszerű jelek. Az időben véletlenszerű jelekre gyakran úgy hivatkozunk, hogy sztochasztikus jelek. Minden osztályban a jelek lehetnek folyamatosak és diszkrét az időben, és rendelkezhetnek folytonos értékű és diszkrét értékű amplitúdóval.

A determinisztikus jel definiálható úgy, hogy egyike az előre meghatározott görbéjű jeleknek időben és térben. A determinisztikus jelek pontos hullámszáma teljes mértékben leírható az idő függvényében, és a jel pontos értéke bármely időpillanatra kiszámítható a függvényi leírásból, és a korábbi értékek is kiszámíthatók ezen a módon. Például egy szinuszos hullám  $x(t)$  modellezhető és pontosan kiszámítható mind egy másodrendű lineáris prediktív modellből, vagy a közismertebb  $x(t) = A \cdot \sin(2\pi ft + \phi)$  egyenlettel.

A véletlenszerű jeleknek nem kiszámítható hullámszámai vannak, ezért nem lehet olyan egyenletet felírni, mellyel kiszámítható lenne a jel pontos jövőbeli értéke a korábbi értékeiből. A legtöbb jel, mint például a beszéd, vagy zaj legalább részben véletlenszerű. A véletlenszerűség elmélete az információ és a zaj elméletével teljesen összekapcsolódott. Valóban, a véletlenszerű jelek feldolgozásakor a munka nagy részét jelenti az információ kinyerése a zajosan vett jelből. Ha egy jel alkalmas információhordozásra, akkor valamilyen fokú véletlenszerűségnek kell benne lennie: a kiszámítható jelek nem hordoznak információt. Ezért a jel véletlenszerű

része lehet a jel információtartalma, vagy zaj, vagy a kettő keveréke, tehát információ és zaj. Bár a véletlenszerű jelek nem teljesen kiszámíthatók, mégis gyakran mutatnak egy sor olyan statisztikailag jellemző értéket, mint például a maximum, a minimum, a középérték, a felező, a szórásnégyzet, és a teljesítményspektrum. A véletlenszerű jel leírható a statisztikája szempontjából, és a legteljesebben a valószínűségi modell szempontjából, melyből az összes statisztikája számítható.

#### 2.2.1.1.1. Sztochasztikus folyamatok

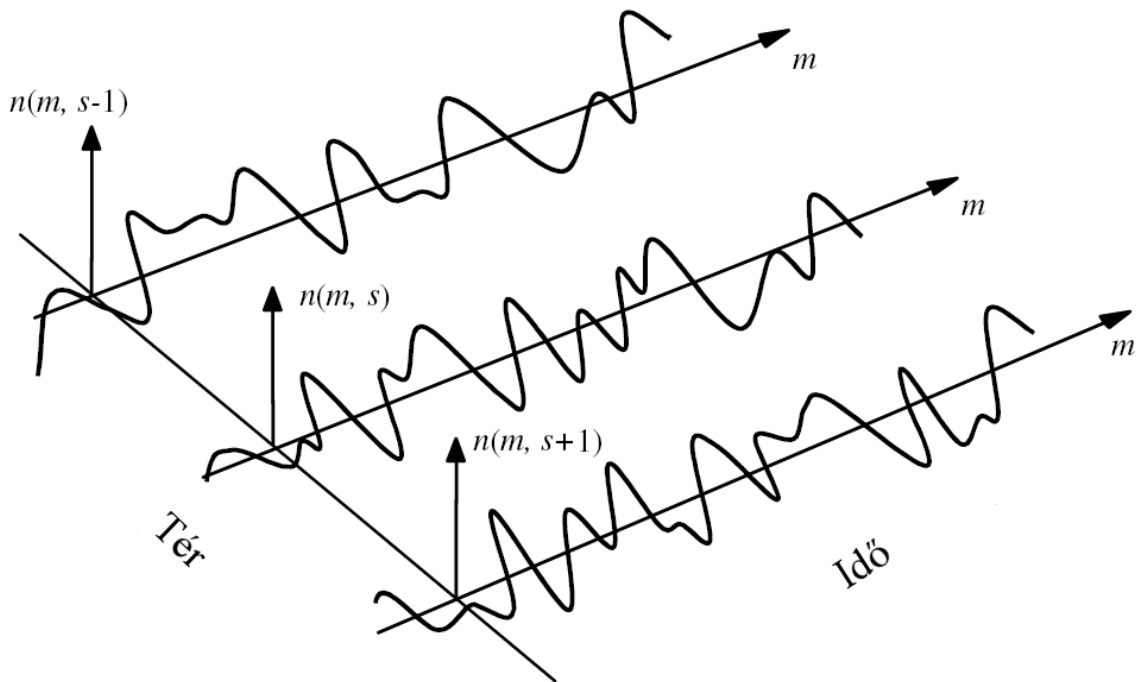
A „sztochasztikus eljárás” szakkifejezést általában a véletlenszerű folyamatok leírására használjuk, melyek folyamatos jeleket állítanak elő, mint például az emberi beszéd, vagy zaj. A jelfeldolgozási terminológiában egy sztochasztikus eljárás egy valószínűségi modell, a véletlenszerű jelek egy osztályára, például Gauss módszer, Markov módszer, Poisson módszer. Egy klasszikus példa a sztochasztikus eljárásra a részecskék úgynevezett Brown mozgása folyadékokban. A részecskék a folyadék belsejében véletlenszerűen mozognak a folyadékrészecskék „bombázása” következtében. Minden részecske véletlenszerű mozgása a sztochasztikus módszer egy egyszerű elképzelése. Minden részecske folyadékban történő mozgása segíti a megértést, vagy a módszer különféle megvalósításainak terét.

#### 2.2.1.1.2. Véletlen folyamatok együttese vagy tere

A véletlenszerű folyamatok összes elképzelésének gyűjteménye a módszerek együttese vagy tere néven ismert. Példaként tekintsünk egy véletlenszerű zajt egy telekommunikációs rendszerben ahogyan azt a 11. ábra is mutatja. A zaj minden telefonvonalon időben véletlenszerűen ingadozik, és jelölhető az  $n(m,s)$  jelöléssel, ahol  $m$  a diszkrét idő mérőszám, és  $s$  jelöli a sorindexet. A zajok gyűjteménye a különböző vonalakon a zaj együttesét (vagy terét) alkotják, melyet a  $N(m)=\{n(m,s)\}$  jelöléssel jelölünk, ahol  $n(m,s)$  az  $N(m)$  zaj realizációját jelöli az  $s$  vonalon.

#### 2.2.1.2. Valószínűségi modellek

A valószínűségi modellek a véletlenszerű jelek legteljesebb matematikai leírását adják. Egy  $m$  rögzített időpillanatban egy véletlenszerű jeltől gyűjtött minta  $\{x(m,s)\}$  egy valószínűségi változó, mely különböző értékeket vesz fel az  $s$  térben. Az alapvető különbség a véletlenszerű változó és a véletlenszerű folyamat között, hogy az utóbbi idősorozatot generál. Ezért a véletlenszerű változók modellezésére használt modellek használhatók a véletlenszerű folyamatokra is.



11. ábra A véletlen zaj három realizációja a zajok terében

A valószínűségi változó tere az összes érték gyűjteménye, vagy eredmény, amit a változó felvehet. A valószínűségi változó tere részekre osztható, néhány kritérium szerint, néhány altérre. Az altér jelértékek egy gyűjteménye, melyek egy közös tulajdonsággal rendelkeznek, mint egy csoportba összezsúfolt minta, vagy a minták gyűjteménye azok amplitúdójával az adott sávon belül. Minden alteret egy eseménynek hívunk, és az esemény bekövetkezésének valószínűsége  $A$ , és  $P(A)$  az a hányados, mely az  $A$  térből származó megfigyelések eredményeinek számát  $N_A$ -t

elosztjuk az összes megfigyelés számával:  $P(A) = \frac{N_A}{\sum_i N_i}$ . Az iménti egyenletből evi-

dens, hogy a kísérletben résztvevő összes lehetséges esemény valószínűségének összege egy.

#### 2.2.1.2.1. Valószínűségi változó eloszlása, eloszlásfüggvény

Legyen adott egy  $(\Omega, \mathbf{A}, P)$  valószínűségi mező, legyen értelmezve egy  $\xi: \Omega \rightarrow \mathbf{R}$  függvény, melyre teljesül, hogy  $\{\omega \mid \xi(\omega) = x\} \in \mathbf{A}$  minden  $x \in \mathbf{R}$  esetén. ( $\xi$  az elemi események halmazát képezi le  $\mathbf{R}$ -be, értékkészlete diszkrét értékekből áll). Ekkor  $\xi$ -t diszkrét valószínűségi változónak nevezzük.  $\xi$  értékkészlete  $\{x_1, x_2, \dots\}$  megszámlálható.  $\xi$  eloszlása:  $\{p_i = P(\xi = x_i) = P(\omega \mid \xi(\omega) = x_i) \mid i=1, 2, \dots\}$  valószínűségek halmaza, annyi elemű, ahány elemű az értékkészlet, elemei mutatják a diszkrét valószínűségi változó egyes értékeinek felvételi valószínűségét. Az értékkészlet elemei nemnegatív számok, melyek összege 1.

Két diszkrét valószínűségi változó ( $\xi$  és  $\eta$ ) *együttes eloszlásán* a következő eloszlást értjük (amennyiben  $\xi$  értékészlete  $\{x_1, x_2, \dots\}$  eloszlása  $\{P(\xi=x_k) \ (k=1, 2, \dots)\}$ ,  $\eta$  értékészlete  $\{y_1, y_2, \dots\}$  eloszlása  $\{P(\eta=y_j) \ (j=1, 2, \dots)\}$ ):  $\{P(\xi=x_k \text{ és } \eta=y_j) \ (k,j=1, 2, \dots)\}$ .

Az együttes eloszlás elemei nemnegatívak, az elemek összege 1. Az együttes eloszlásban szereplő események teljes eseményrendszert alkotnak, ezért az előző állítás igaz.

A peremeloszlás két diszkrét valószínűségi változó együttes eloszlásából meghatározható az alkotó diszkrét valószínűségi változók eloszlása. Valamelyik diszkrét valószínűségi változó eloszlása szerint összegezve az együttes eloszlást, a másik diszkrét valószínűségi változó eloszlásához jutunk.

Diszkrét valószínűségi *változók függetlenek*, ha minden  $j = 1, 2, \dots$  és minden  $k = 1, 2, \dots$  esetén  $P(\xi=x_k \text{ és } \eta=y_j)=P(\xi=x_k)P(\eta=y_j)$ , azaz együttes eloszlásuk minden eleme szétbomlik az eloszlások megfelelő elemeinek szorzatára. Két diszkrét valószínűségi változó eloszlása pontosan akkor határozza meg együttes eloszlásukat, ha függetlenek. A konstans értékű valószínűségi változó minden más diszkrét valószínűségi változótól független.

Adott valószínűségi mezőben lévő  $B_i \ (i=1, 2, \dots)$  teljes eseményrendszer esetén ha ezek között nincs 0 valószínűségű esemény, akkor tetszőleges nem 0 valószínűségű  $A$  eseményre  $P(A | B) = \frac{P(A | B_i) \cdot P(B_i)}{\sum_i (P(A | B_i) \cdot P(B_i))}$ . Ez Bayes tétele.

Egy  $\xi$  valószínűségi változó *eloszlásfüggvénye*:  $F: \mathbf{R} \rightarrow [0, 1]$ ,  $F(x)=P(\xi < x)$ . Valószínűségi változó eloszlásfüggvénye balról folytonos, monoton nemcsökkenő,  $-\infty$ -ben 0-hoz tart,  $+\infty$ -ben 1-hez tart. Bármely, a valószínűségi változó eloszlásfüggvényeire vonatkozó fenti feltételeket kielégítő függvényhez létezik olyan valószínűségi változó, melynek eloszlásfüggvénye az adott függvény. A diszkrét valószínűségi változónak is létezik eloszlásfüggvénye, ez „lépcsős” alakú.  $\xi$  valószínűségi változó esetén  $P(a \leq \xi < b) = P(\xi < b) - P(\xi < a) = F(b) - F(a)$ , ha  $a < b$ .  $\xi$  és  $\eta$  *valószínűségi változók együttes eloszlásfüggvénye*  $F(x,y)=P(\xi < x \text{ és } \eta < y) \ (x,y \in \mathbf{R})$ .  $\xi$  és  $\eta$  *valószínűségi változók függetlenek*, ha együttes eloszlásfüggvényük megegyezik eloszlásfüggvényeik szorzatával. Valószínűségi változók együttes eloszlásfüggvénye mindkét változójában monoton nemcsökkenő, balról folytonos,  $\lim_{x \rightarrow -\infty} F(x,y) = \lim_{y \rightarrow -\infty} F(x,y) = 0$ ,  $\lim_{y \rightarrow \infty, x \rightarrow \infty} F(x,y) = 1$ ; ha  $a_1 < b_1$  és  $a_2 < b_2$ , akkor  $F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) - F(a_1, a_2) \geq 0$ . Bármely, a valószínűségi változók együttes eloszlásfüggvényére vonatkozó ezen feltételeket kielégítő függvényhez létezik két olyan valószínűségi változó, melyek-

nek együttes eloszlásfüggvénye az adott függvény. Az utolsó feltétel  $P(\xi \in [a_1, b_1] \text{ és } \eta \in [a_2, b_2]) \geq 0$ -t jelenti.

### 2.2.1.2.2. Sűrűségfüggvény

Egy adott eseménytéren értelmezett  $\xi: \Omega \rightarrow \mathbf{R}$  függvény *valószínűségi változó*, ha  $\{\omega \mid \xi(\omega) < x\} \in \mathbf{A}$  bármely  $x \in \mathbf{R}$  esetén. A diszkrét valószínűségi változó teljesíti az iménti definíció követelményeit. A definíció a diszkrét valószínűségi változók általánosítása: folytonos esetben  $P(\xi=x)=0$  minden  $x \in \mathbf{R}$  esetén. Valószínűségi változót (ha nem hangsúlyozottan diszkrét) a definíció szerinti értelemben kell érteni.

Egy  $\xi$  valószínűségi változó *folytonos*, ha eloszlásfüggvénye abszolút folytonos. Folytonos valószínűségi változó esetén  $P(\xi=x)=0$  bármely  $x \in \mathbf{R}$  esetén.

Belátható, hogy  $0 \leq P(\xi=x) = \lim_{\Delta x \rightarrow 0} P(x \leq \xi < x + \Delta x) = \lim_{\Delta x \rightarrow 0} (F(x + \Delta x) - F(x)) \leq \lim_{\Delta x \rightarrow 0} ((F(x) + K\Delta x) - F(x)) = \lim_{\Delta x \rightarrow 0} (K\Delta x) = 0$ .

Bármely  $\xi$  valószínűségi változó esetén bármely  $x \in \mathbf{R}$ -re  $P(a \leq \xi \leq b) = F(b+0) - F(a)$ ,  $P(a < \xi < b) = F(b) - F(a-0)$ ,  $P(a < \xi \leq b) = F(b+0) - F(a-0)$  teljesül. Folytonos esetben ezek mind egyenlők  $P(a \leq \xi < b)$ -vel. Folytonos  $\xi$  valószínűségi változó *sűrűségfüggvénye*

$f: \mathbf{R} \rightarrow \mathbf{R}$ , ha  $\int_{-\infty}^x f(t) dt = F(x)$  teljesül minden  $x \in \mathbf{R}$  esetén. Folytonos valószínűségi

változó sűrűségfüggvényére és eloszlásfüggvényére teljesül az  $F'(x) = f(x)$  összefüggés. Diszkrét valószínűségi változónak nincs sűrűségfüggvénye.

A sűrűségfüggvény nemnegatív, az egész számegegyenesen vett integrálja 1. Bármely, a valószínűségi változó sűrűségfüggvényére vonatkozó, ezen feltételeket kielégítő függvényhez létezik olyan valószínűségi változó, melynek sűrűségfüggvénye az adott függvény. Diszkrét esetben hasonló feltétel a  $p_i \geq 0$ ,  $\sum p_i = 1$  tulajdonságok: az eloszlás játssza a sűrűségfüggvény szerepét is.

$\xi$  és  $\eta$  abszolút folytonos *valószínűségi változók együttes sűrűségfüggvénye*

$f(u, v)$ , ha rá teljesül az  $\int_{-\infty}^x \int_{-\infty}^y f(u, v) dv du = F_{\xi, \eta}(x, y)$  egyenlőség.

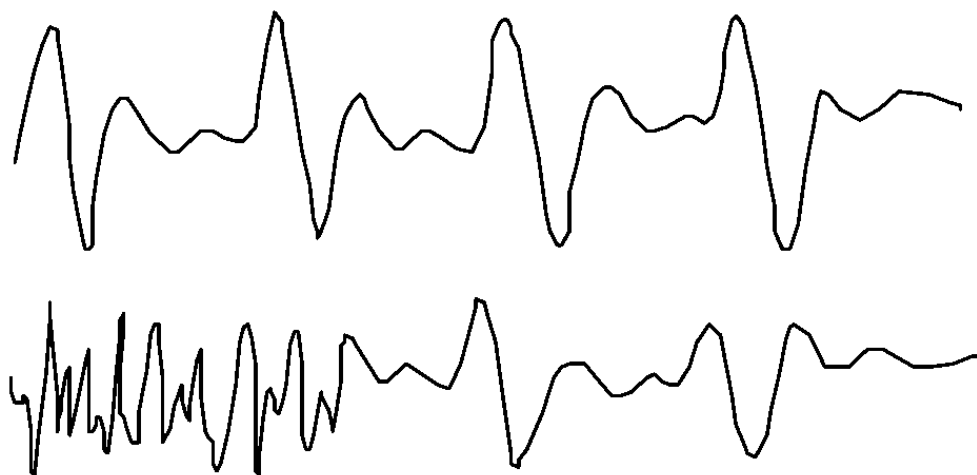
Az együttes sűrűségfüggvény megegyezik az együttes eloszlásfüggvény második parciális deriváltjával (először az első, majd a második változó szerint deriválva). Abszolút folytonos  $\xi$  és  $\eta$  valószínűségi változók esetén függetlenségük egyenértékű  $f_{\xi, \eta}(u, v) = f_{\xi}(u)f_{\eta}(v)$  teljesülésével. Az együttes sűrűségfüggvény nemnegatív, az egész valós számsíkon vett integrálja 1. Az együttes sűrűségfüggvényből meghatározható az alkotó valószínűségi változók sűrűségfüggvénye. Az együttes sűrűség-



függvényt az egyik változó szerint az egész valós számegyenesen integrálva a másik valószínűségi változó sűrűségfüggvényét kapjuk.

### 2.2.1.3. Állandó és nem állandó véletlen folyamatok

Bár a jel  $x(m)$ -mel jelölt amplitúdója  $m$  periódussal hullámzik, a jel által generált folyamat jellemzője lehet időben állandó vagy időben változó. Nem állandó folyamat például a beszéd, melynek hangossága és spektrális jellemzői, összetevői folyamatosan változnak, ahogyan a beszélő különféle hangokat képez. A folyamat állandónak tekinthető, ha a valószínűségi modellbeli jellemzői idő invariánsak, ellenkező esetben nem állandó – mindkettőt a 12. ábra szemlélteti. Az állandó tulajdonság magába foglalja, hogy a jel összes paramétere, úgy mint középérték, variancia, energiaspektrum összetétele, és a folyamat magasabb rendű momentumai idő invariánsak. A gyakorlatban az állandóságnak különféle fokai ismeretesek. Van olyan, hogy a statisztikák egy halmaza állandó, a többi pedig nem. Például lehet egy véletlen folyamatnak idő invariáns középértéke, de az idő függvényében változó energiája.



12. ábra Kvázistacionárius és nemstacionárius beszédsszegmensek

### 2.2.1.3.1. Szigorú értelemben vett állandó folyamatok

Egy  $X(m)$  véletlenszerű folyamat szigorú értelemben véve stacionárius, ha minden megoszlása, és statisztikai paramétere idő invariáns. A szigorú értelemben vett állandóság magában foglalja, hogy az  $n$ -edik rendű eloszlása transláció invariáns minden  $n = 1, 2, \dots$ , vagyis:

$$P[x(m_1) \leq x_1, x(m_2) \leq x_2, \dots, x(m_n) \leq x_n] = P[x(m_1 + \tau) \leq x_1, x(m_2 + \tau) \leq x_2, \dots, x(m_n + \tau) \leq x_n]$$

Az egyenletből a szigorú értelemben vett állandó folyamat statisztikája beleértve a középértéket, a korrelációt és az energiaspektrumot, idő invariáns lesz, ezért:  $E[x(m)] = \mu_x$ ,  $E[x(m)x(m+k)] = r_{xx}(k)$  és  $E[|X(f, m)|^2] = E[|X(f)|^2] = P_{xx}(f)$ , ahol  $\mu_{xx}$ ,  $r_{xx}(m)$  és  $P_{xx}(f)$  rendre az  $x(m)$  jel középértéke, az autokorrelációja és az energiaspektruma, és  $X(f, m)$  jelöli  $x(m)$  frekvencia-idő spektrumát.

### 2.2.1.3.2. Nem szigorú értelemben vett állandó folyamatok

Míg a szigorú értelemben vett állandó (stacionárius) folyamatok esetében a minden statisztikának idő-invariánsnak kell lennie, a nem szigorú értelemben vett stacionárius folyamatokra nem vonatkozik az összes korlátozás. Egy ilyen folyamat középértéke és az autokorrelációs függvénye idő invariáns:  $E[x(m)] = \mu_x$ ,  $E[x(m)x(m+k)] = r_{xx}(k)$ . A két folyamat definíciójából kiderül, hogy minden szigorú értelemben vett stacionárius folyamat nem szigorú értelemben is az, viszont ennek a fordítottja nem igaz.

### 2.2.1.3.3. Nem stacionárius folyamatok

Egy véletlenszerű folyamat nem stacionárius, ha eloszlása vagy statisztikája időben változik. A legtöbb sztochasztikus folyamat, mint például a videojelek, audiojelek, gazdasági adatok, meteorológiai adatok, stb., nem stacionárius, mert egy-egy olyan rendszer generálja őket, melynek környezeti paraméterei időben változnak. Például a beszéd nem stacionárius, mert egy időben változó artikulációs rendszer állítja elő. A beszéd hangerőssége, frekvencia-összetétele időben változó, mely változás néha teljesen váratlan. Az időben változó folyamatok modellezhetők stacionárius véletlenszerű folyamatok egy kombinációjával.

### 2.2.1.4. Véletlen folyamatok várható értéke

A várható érték központi szerepet játszik a jelek modellezésében és feldolgozásában. Továbbá a véletlenszerű folyamatok valószínűségi modelljei általában a várható érték függvényeként kerülnek kifejezésre. Például a normális eloszlás el-

oszlásfüggvénye a folyamat középértékét és kovarianciáját tekintve egy exponenciális függvényként definiálható, a Poisson eloszlás eloszlásfüggvénye pedig a folyamat középértékével definiálható. A jelfeldolgozási alkalmazásokban mindig megvan a jel megfelelő statisztikai modellje – például a normális eloszlás eloszlásfüggvénye –, és a modell kiegészítésére szükség van a várható paraméterek értékére. Azonkívül sok jelfeldolgozási algoritmusban, mint például a spektrális kivonásban a zajcsökkentésre, vagy a lineáris előrejelzésben, lényegében amire szükségünk van az a jel középértékének vagy a korrelációs függvényének kiszámítása. Egy véletlenszerű folyamat  $h(X(m_1), X(m_2), \dots, X(m_M))$  függvényének várható értékének kiszámítása a következőképpen definiálható:

$$E[h(X(m_1), \dots, X(m_M))] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} h(x_1, \dots, x_M) f_{X(m_1) \dots X(m_M)}(x_1, \dots, x_M) dx_1 \dots dx_M$$

A legfontosabb és széles körben használt várható értékek a középérték, a korreláció, a kovariancia és a energiaspektrum.

#### 2.2.1.4.1. A középérték

A egy jelsorozat középértéke nagyon fontos szerepet játszik a jelfeldolgozásban és paraméterkiszámításban a zajos jelek esetében. Például egy zajos jel esetében az optimális lineáris kiszámítási módszer az interpoláció a középérték és a megfigyelt jel zajjal terhelt értéke között. Egy  $[X(m_1), \dots, X(m_M)]$  véletlenszerű vektor középértéke alatt értjük a teljes jelfolyam alatt mutatott átlagértéket, mely a következőképpen definiálható:

$$E[X(m_1), \dots, X(m_M)] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (x_1, \dots, x_M) f_{X(m_1) \dots X(m_M)}(x_1, \dots, x_M) dx_1 \dots dx_M$$

#### 2.2.1.4.2. Autokorreláció

A korrelációs függvény és annak Fourier transzformációja, az energiaspektrum sűrűsége minták és struktúrák modellezésére és felismerésére használatos a jelfeldolgozások során. A kölcsönösségi együtthatók központi szerepet játszanak a jelfeldolgozásban, a telekommunikációs rendszerekben, beleértve a prediktív kódolókat, kiegyenlítőket, digitális dekódereket, késleltetés kiszámítókat, osztályozókat és jelvisszaállító rendszereket. Egy  $X(m)$  véletlenszerű folyamat autokorrelációs függvénye, melyet  $r_{xx}(m_1, m_2)$ -vel jelölünk, a következőképpen definiálható:

$$r_{xx}(m_1, m_2) = E[x(m_1)x(m_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(m_1)x(m_2) f_{X(m_1), X(m_2)}(x(m_1), x(m_2)) dx(m_1) dx(m_2)$$

Az  $r_{xx}(m_1, m_2)$ -vel jelölt autokorrelációs függvény az  $X$  folyamat végeredményének  $m_1$  és  $m_2$  időpillanatbeli hasonlóságának mértékét vagy a kölcsönösségi relációját mutatja meg. Ha a véletlenszerű folyamat végeredménye az  $m_1$  és az  $m_2$  időpillanatbeli értéke nincs hordoz relációt, akkor  $X(m_1)$  és  $X(m_2)$  függetlennek vagy korrelálatlannak mondható, tehát  $r_{xx}(m_1, m_2) = 0$ . A nem szigorú értelemben vett stacionárius folyamatok esetében az autokorrelációs függvény idő invariáns és a két időpillanat között eltelt idő hosszának függvénye, vagyis  $m = m_1 - m_2$ :  $r_{xx}(m_1 + \tau, m_2 + \tau) = r_{xx}(m_1, m_2) = r_{xx}(m_1 - m_2) = r_{xx}(m)$ . Egy valós értékű nem szigorú értelemben vett stacionárius folyamat autokorrelációs függvénye szimmetrikus a következő tulajdonságokkal:  $r_{xx}(-m) = r_{xx}(m)$  és  $r_{xx}(m) \leq r_{xx}(0)$ . Ez esetben az  $r_{xx}(0)$  jelenti a nulla középértékű jel energiáját.

#### 2.2.1.4.3. Autokovariancia

Az autokovariancia függvény egy  $X(m)$  véletlenszerű folyamat  $c_{xx}(m_1, m_2)$  jelöléssel ellátott szórásának vagy szóródásának mérésére szolgál a folyamat középértéke körül és a következőképpen definiálható:

$$c_{xx}(m_1, m_2) = E[(x(m_1) - \mu_x(m_1))(x(m_2) - \mu_x(m_2))] = r_{xx}(m_1, m_2) - \mu_x(m_1)\mu_x(m_2),$$

ahol a  $\mu_x(m)$  jelenti az  $X(m)$  véletlenszerű folyamat középértékét. Nulla középértékű folyamatok esetében az autokovariancia és az autokorreláció függvények azonosak. A  $c_{xx}(m_1, m_2)$  a folyamat varianciáját (szórásnégyzetét) jelenti. Stacionárius folyamatok esetében az autokovariancia függvény az előző egyenletből a következőképpen alakul:  $c_{xx}(m_1, m_2) = c_{xx}(m_1 - m_2) = r_{xx}(m_1 - m_2) - \mu_x^2$ .

#### 2.2.1.4.4. Spektrális energiasűrűség

Egy véletlenszerű folyamat energiájának spektrális sűrűségfüggvénye, más néven energiaspektruma megadja a jel egyes frekvenciákhoz tartozó energiáinak eloszlását a teljes spektrumon. A nem szigorú értelemben vett stacionárius folyamat  $X(m)$  energiaspektruma a Wiener–Khinchin tétel alapján úgy definiálható, hogy az autokorrelációs függvény Fourier transzformáltja és a következőképpen írható fel:  $P_{XX}(f) = E[X(f)X^*(f)] = \sum_{m=-\infty}^{\infty} r_{xx}(m)e^{-j2\pi fm}$ , ahol  $r_{xx}(m)$  és  $P_{XX}(f)$  rendre az  $x(m)$  autokorrelációs függvénye és energiaspektruma, és  $f$  pedig a frekvenciaváltozó. Egy valós értékű stacionárius folyamat autokorrelációs függvénye szimmetrikus, és az energiaspektruma a következőképpen írható fel:

$$P_{XX}(f) = r_{xx}(0) + \sum_{m=1}^{\infty} 2r_{xx}(m) \cos(2\pi fm)$$

A spektrális sűrűségfüggvény egy nemnegatív valós értékészletű felvevő függvény, megadása pedig watt per herzben  $\left[\frac{\text{W}}{\text{Hz}}\right]$  történik. Az autokorrelációs so-

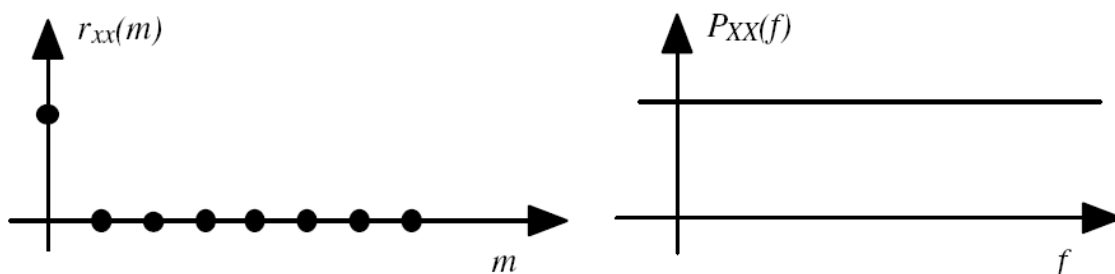
rozat a véletlenszerű folyamat energiaspektrumából az iménti egyenlet felhasználásával kapható meg az inverz Fourier transzformáció alkalmazásával a követ-

zőképpen:  $r_{xx}(m) = \int_{-\frac{1}{2}}^{\frac{1}{2}} P_{XX}(f) e^{j2\pi fm} df$ . Az autokorreláció és az energiaspektrum a

folyamat másodrendű statisztikai, rendre időben és frekvenciatartományban.

#### Példa – fehér zaj

A következő ábra a fehér zaj energiaspektrumát és autokorrelációját mutatja.



13. ábra A fehér zaj autokorrelációja és energiaspektruma

Egy olyan zajt, amely korrelálatlan, egymástól független mintákból áll, fehér zajnak nevezünk. Egy stacionárius fehér zaj  $n(m)$  autokorrelációja definiálható:

$$r_{nn}(k) = E[n(m)n(m+k)] = \begin{cases} \text{Zajenergia} & k = 0 \\ 0 & k \neq 0 \end{cases}. \text{ Ez az egyenlet egy korrelálatlan fehér}$$

zaj definíciója szerinti matematikai állítás. A frekvenciatartományra ekvivalens leírás adható meg, az  $r_{nn}(k)$  Fourier transzformációjával a következőképpen:

$$P_{NN}(f) = \sum_{k=-\infty}^{\infty} r_{nn}(k) e^{-j2\pi fk} = r_{nn}(0) = \text{zajenergia}. \text{ A stacionárius fehér zaj energiaspekt-}$$

ruma azonosan terjed szét a sáv teljes szélességében és időben. A fehér zaj az egyik legnehezebben eltávolítható zajforma, mert nem határozható meg a szerkezete sem időben, sem a frekvenciasávok szerint.

#### Példa – impulzív zaj

Az impulzív zaj egy véletlenszerű kétállapotú („be/ki”) impulzussorozatból áll, véletlenszerű amplitúdóval és időben véletlenszerű előfordulással. Egy véletlenszerű impulzív zaj sorozat  $n_i(m)$  egy kétállapotú véletlenszerű amplitúdó modulált sorozatként modellezhető, a következő szerint:  $n_i(m) = n(m)b(m)$ , ahol  $b(m)$  egy

kétállapotú sorozat, mely egy impulzus jelenlétét vagy hiányát jelzi, és  $n(m)$  pedig egy véletlenszerű zaj. Tegyük fel, hogy az impulzív zaj egy korrelálatlan folyamat, így az impulzív zaj, mint egy kétállapotú folyamat autokorrelációja a következőképpen definiálható:  $r_m(k, m) = E[n_i(m)n_i(m+k)] = \sigma_n^2 \delta(k)b(m)$ , ahol  $\sigma_n^2$  a zaj varianciája. Ebben az egyenletben az autokorrelációs függvény egy kétállapotú függvényként lett kifejezve, mely az impulzív zaj  $m$  időpillanatbeli be/ki állapotától függ. Az impulzív zaj sorozat energiaspektruma az imént leírt autokorrelációs függvény Fourier transzformáltjaként kapható meg a következő összefüggéssel:  $P_{NN}(f, m) = \sigma_n^2 b(m)$ .

#### 2.2.1.4.5. Két véletlenszerű folyamat együttes statisztikája

A jelfeldolgozás során gyakran találkozunk azzal a problémával, például egy több érzékelőből álló rendszer esetében, hogy nem csak egy véletlenszerű folyamat van jelen. Az együttes statisztika és együttes eloszlás leírja a két vagy több folyamat kapcsolatát egymással. Két diszkrét időbeosztással mintavételezett véletlenszerű folyamat  $x(m)$  és  $y(n)$  együttes eloszlásfüggvénye a következő jelöléssel adható meg:  $f_{X(m_1)\dots X(m_M), Y(n_1)\dots Y(n_N)}(x_1, \dots, x_M, y_1, \dots, y_N)$ . Ha a két véletlenszerű folyamat  $X(m)$  és  $Y(n)$  korrelálatlanok, az együttes eloszlásfüggvényük a két eloszlásfüggvény szorzataként adható meg a következő felírással:

$$f_{X(m_1)\dots X(m_M), Y(n_1)\dots Y(n_N)}(x_1, \dots, x_M, y_1, \dots, y_N) = f_{X(m_1)\dots X(m_M)}(x_1, \dots, x_M) f_{Y(n_1)\dots Y(n_N)}(y_1, \dots, y_N)$$

#### 2.2.1.4.6. Kereszt-korreláció és kereszt-kovariancia

Két véletlenszerű folyamat kereszt-korrelációja az alábbiak szerint adható

$$\text{meg: } r_{xy}(m_1, m_2) = E[(x(m_1)y(m_2))] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(m_1)y(m_2)f_{X(m_1)Y(m_2)}(x(m_1), y(m_2))dx(m_1)dy(m_2)$$

Egy  $r_{xy}(m_1, m_2)$  nem szigorú értelemben vett stacionárius folyamat esetében a korrelációs függvény csak az  $m_1$  és  $m_2$  időpillanatok különbségétől függ, vagyis  $m = m_1 - m_2$ , tehát:  $r_{xy}(m_1 + \tau, m_2 + \tau) = r_{xy}(m_1, m_2) = r_{xy}(m_1 - m_2) = r_{xy}(m)$ .

A kereszt-kovariancia függvény a következőképpen definiálható:  $c_{xy}(m_1, m_2) = E[(x(m_1) - \mu_x(m_1))(y(m_2) - \mu_y(m_2))] = r_{xy}(m_1, m_2) - \mu_x(m_1)\mu_y(m_2)$ . Nulla középértékkel rendelkező folyamatok esetében a kereszt-korreláció és a kereszt-kovariancia függvények azonosak. Nem szigorú értelemben vett stacionárius folyamatok esetében a kereszt-kovariancia függvény az előző egyenlet alapján:  $c_{xy}(m_1, m_2) = c_{xy}(m_1 - m_2) = r_{xy}(m_1 - m_2) - \mu_x\mu_y$ .

## 2.2.1.4.7. Kereszt-energiaspektrum eloszlás és koherencia

Két véletlenszerű folyamat  $X(m)$  és  $Y(m)$  kereszt-energiaspektrum eloszlása a két folyamat kereszt-korrelációs függvényének Fourier transzformációjával definiálható:

$P_{XY}(f) = E[X(f)Y^*(f)] = \sum_{m=-\infty}^{\infty} r_{xy}(m)e^{-j2\pi fm}$ . Ahogyan a kereszt-korreláció a

kereszt-energiaspektrum eloszlása is a két jel hasonlóságának mértékét mutatja, vagy a két folyamat spektrális koherenciáját, mely a kereszt-energiaspektrum eloszlásának normalizált formulája és a következőképpen definiált:

$C_{XY}(f) = \frac{P_{XY}(f)}{\sqrt{P_{XX}(f)P_{YY}(f)}}$ . A koherencia függvény olyan alkalmazásokban haszná-

latos, mint az időkésettetés kiszámítása és a jel-zaj arány meghatározása (mérése).

## 2.2.2. A rejtett Markov modell

A rejtett Markov modelleket a véletlenszerű nem stacionárius jelek statisztikai modellezésére használhatjuk hatékonyan. Az ilyen jelek csoportjába tartozik az emberi beszéd vagy az időben változó zaj. A HMM egy Markov láncot ad meg, mely az időbeli változást vagy a terek változását állapotfüggő stacionárius alfolyamatokra bontja. Egy HMM lényegében egy véges állapothalmazzal rendelkező Bayes folyamat, Markov előzménnyel, az egyes állapotok között átmenetek modellezésére, és egy halmaz valószínűségi eloszlásfüggvény a jel véletlenszerű variációinak modellezésére minden állapoton belül.

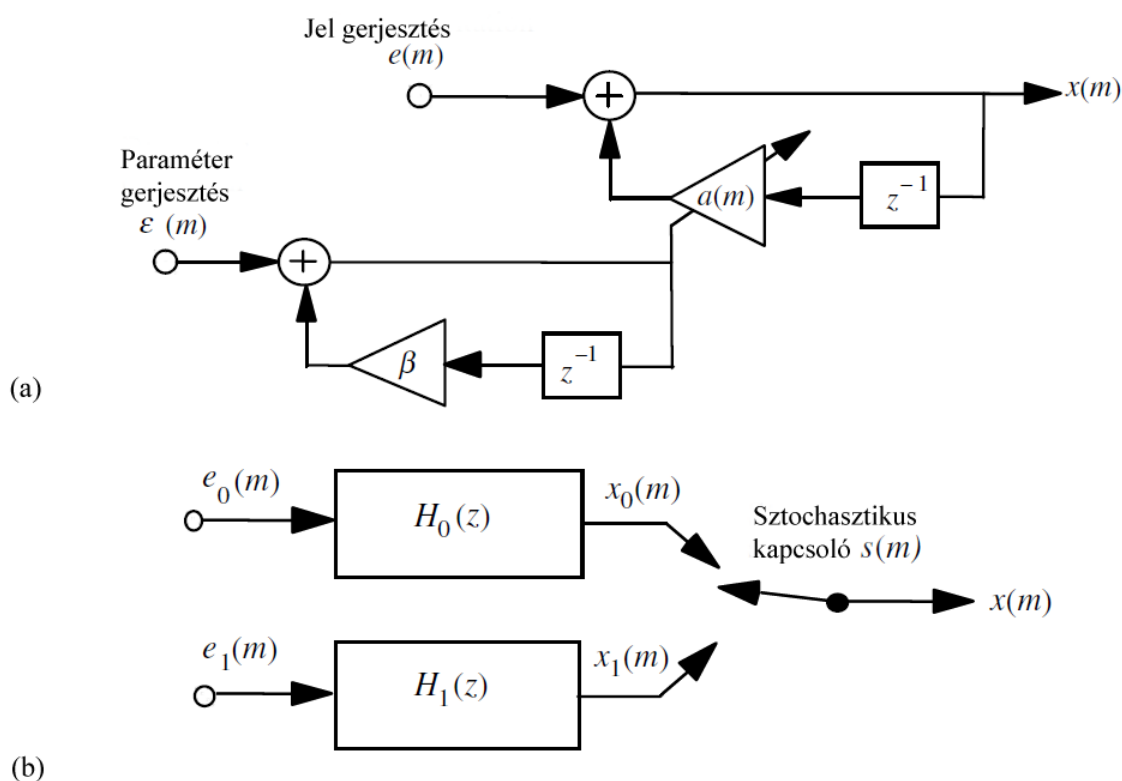
2.2.2.1. Nem stacionárius folyamatok statisztikai modellje

Egy nem stacionárius folyamat definiálható oly módon, hogy egy olyan folyamat, melynek statisztikai paraméterei az idővel változnak. A legtöbb „természetesen előállított” jel, mint például a hangjelek, képek, biomedikai és szeizmikus jelek nem stacionárius jelek, mivel az azokat előállító rendszerek paraméterei és a környezet, melyben terjednek változnak az idő múlásával.

A nem stacionárius jelek egy kétrétegű sztochasztikus folyamatként modellezhetők, egy rejtett folyamattal az észlelhető jel statisztikai paramétereinek időbeli változásainak kontrollálására. Általában a nem stacionárius folyamatok két alapvető nagy csoportba sorolhatók:

- a.) folytonosan változó állapotú folyamatok
- b.) véges sok állapotú folyamatok

A folytonosan változó állapotú folyamatok definíciós szerint az időben folytonosan változó jelek statisztikájának alapjául szolgálnak. Ilyenek például az audiojelek, a beszéd és a zene, melyeknek az energiája és spektrális összetétele az idővel változik. Véges sok állapottal rendelkező folyamat az olyan statisztikai jellemzőkkel rendelkező folyamat, mely képes váltani a véges sok stacionárius vagy nem stacionárius állapot között. Például az impulzív zaj egy ilyen kétállapotú folyamat. Folytonosan változó folyamatok közelíthetők hozzájuk illő véges sok állapotú folyamattal. A következő ábra egy nem stacionárius autoregresszív (AR) folyamatot mutat be. Ez a folyamat a jel paramétereire egy rejtett stacionárius AR modell és magára a jelre egy időben változó AR modell kombinációját alkalmazza. A



14. ábra Folytonosan változó (a) és kétállapotú (b) AR folyamat

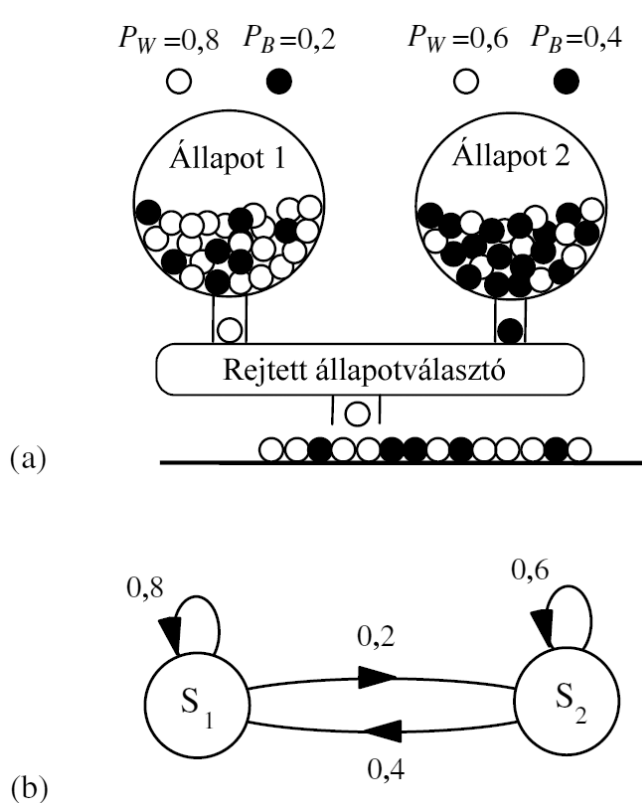
rejtett modell vezérli a nem stacionárius AR modell időben változó paramétereit. Ehhez a modellhez a megfigyelt jel egyenlete és a paraméter állapotegyenlete a következő módon írható fel: a megfigyelt jel  $x(m) = a(m)x(m-1) + e(m)$  és a rejtett állapot  $a(m) = \beta a(m-1) + \varepsilon(m)$ , ahol  $a(m)$  a megfigyelt AR folyamat időben változó együtthatója és  $\beta$  pedig a rejtett állapotirányító folyamat együtthatója. A 14. (b) ábra egy egyszerű példát mutat a véges sok állapotú nem stacionárius folyamat kétállapotú autoregresszív megvalósítására, amelyben egy közvetlen kapcsoló választja ki minden időpillanatban a két AR modell közül az egyiket, melyet a



kimenetre csatol. Ehhez a modellhez a kimeneti jel  $x(m)$  a következő egyenlettel fejthető ki:  $x(m) = \bar{s}(m)x_0(m) + s(m)x_1(m)$ , ahol a kétállású kapcsoló  $s(m)$  kiválasztja a folyamat állapotát  $m$  időpillanatban, és  $\bar{s}(m)$  jelöli  $s(m)$  ellentettjét.

### 2.2.2.2. Rejtett Markov modellek

A rejtett Markov modell (HMM) egy kétrétegű véges sok állapotú folyamat, egy rejtett Markov folyamattal, mely irányítja az állapotok kiválasztását a figyelembe veendő folyamatban. Mint egy egyszerű illusztrációt figyeljük meg a következő ábrán látható kétállapotú Markov folyamatot, mely két tárolót mutat különböző keverékű fehér és fekete színű golyókkal. A fekete és fehér golyók előfordulásának valószínűségét jelöljük rendre  $P_B$  és  $P_W$  jelölésekkel, ahogyan az ábra mutatja. Tegyük fel, hogy folytonos idő intervallumban egy rejtett kiválasztó folyamat kivá-



15. ábra Kétrétegű véletlenszerű folyamat (a), és rejtett Markov modellje (b)

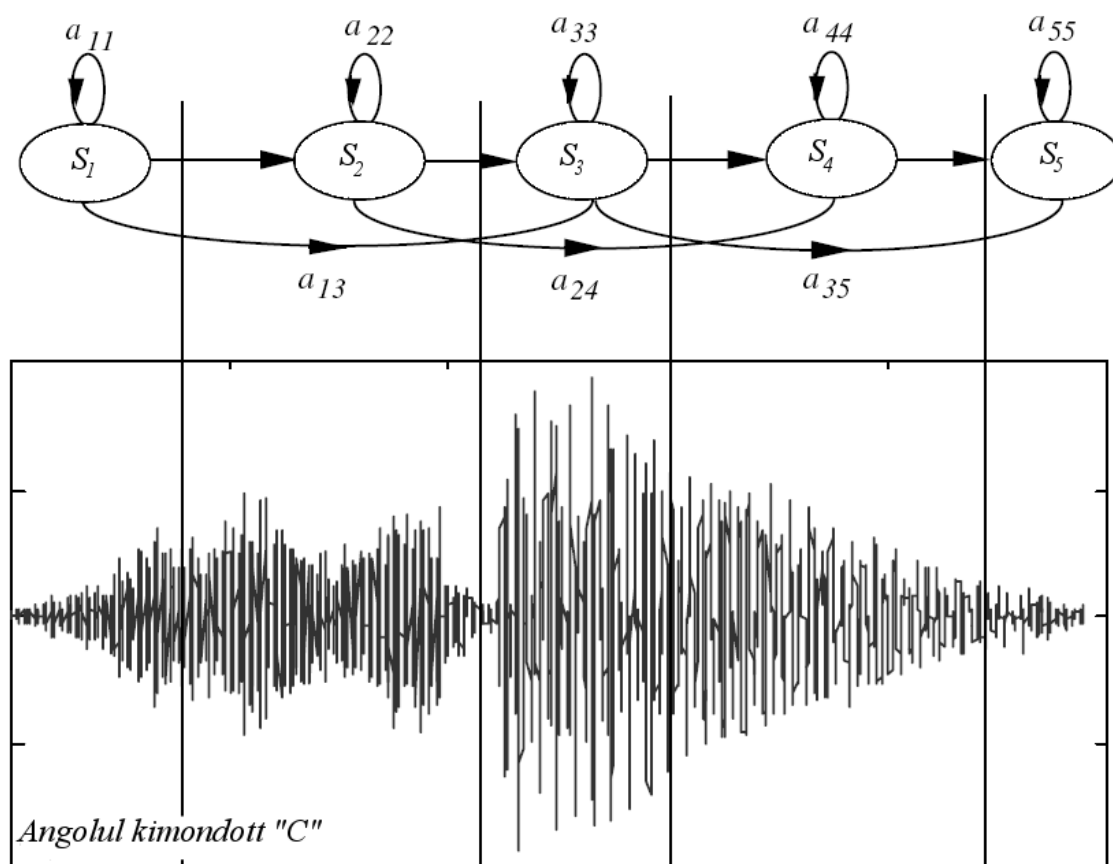
lasztja az egyik tartályt, hogy kiejtsen egy golyót. A golyót kiejtése után pótoljuk, tehát a fehér és fekete golyók számára nincs hatása a műveletnek. Minden tárolót úgy tekinthetünk, hogy az egy alapvető állapota a kimeneti folyamatnak. Most tegyük fel, hogy a rejtett tároló-kiválasztót a következő szabály irányítja: bármely időpillanatban, ha a kiválasztott tárolóból fehér golyót kapunk, akkor nem választunk másik tárolót, ha feketét akkor a másik tárolót választjuk. Ez egy példa a Markov folyamatra, mert a folyamat következő állapota függ a jelenlegi állapottól, ahogyan azt az ábra kétállapotú modellje is mutatja. Ebben a modellben a megfigyelhető végeredmény nem jelzi egyértelműen az alapul vett állapotot, mivel mindkét állapot (tároló) alkalmas mindegyik színű (fekete és fehér) golyó kibocsátására.

Általában a Markov modellnek  $N$  állapota van, és mindegyik állapot más szegmensét modellezi a vizsgált jelnek. A rejtett Markov modell alkalmazható egy

időben változó véletlenszerű folyamat modellezésére, mint egy  $N$  elemű stacionárius vagy kvázi-stacionárius elemi alfolyamat valószínűségi Markov lánc.

### 2.2.2.2.1. A rejtett Markov modellek fizikai értelmezése

A modell fizikai értelmezéséhez nézzük meg a következő, 16. ábrát. Az ábra balról jobbra mutatja egy angolul kiejtett „C”<sup>1</sup> hang modelljét és vele együtt a hang hullámformájának térképét is. Általában kétféle eltérés van a beszéd és az egyéb sztochasztikus jelek között: a spektrális összeállítás változása, és az időbeosztás vagy artikuláció változása. Egy rejtett Markov modellben ezek a változások állapot megjegyzéssel és az egyes állapotok közötti átmenetek valószínűségével modellezhetők. Használható módszer a rejtett Markov modell interpretálására és alkalmazására, hogy a modell minden állapotát úgy tekintjük, hogy a sztochasztikus folyamat egy szegmensét modellezi. Az ábrán az  $S_1$  állapot modellezi az első szeg-



16. ábra Ötállapotú balról–jobbra HMM beszédmodell

<sup>1</sup> Az angol abc 3. betűjének, a c-nek kiejtése a magyartól jelentősen eltér, fonetikusan a következőképpen írható le: [si:]

menet, a  $S_2$  a másodikat és így tovább. Minden állapotnak kell lennie egy olyan mechanizmusának, mely összeilleszti a szegmensek véletlenszerű variációinak különböző realizációit, melyet modelleznek. Az állapotátmenetek valószínűségei egy mechanizmust biztosítanak a különböző állapotok kapcsolatára, és a variációk modellezésére az időskála jelek alatti időtartamára minden állapotban. Például ha a beszéd egy szegmensének vége kinyúlt, köszönhetően a lassú artikulációnak, hozzáigazítható lesz több önmagába visszatérő átmenettel az állapothoz, mely a szegmenst modellezi. Ellenben ha a szegmens elhallgatott a gyors beszédnek köszönhetően, akkor a következő állapotra ugrással illeszt ebben az esetben. Az állapot tapasztalati eloszlásfüggvényei modellezi a jel állapotokhoz rendelt szegmenseinek spektrális összeállításának valószínűségi eloszlását.

#### 2.2.2.2.2. A rejtett Markov modell paraméterei

A rejtett Markov modellnek a következő paraméterei vannak:

- Az állapotok száma  $N$ . Ezt általában beállítjuk az egymástól különböző vagy elemi sztochasztikus események számára egy jelfolyamatban. Például egy impulzív zajhoz hasonló kétállapotú folyamat modellezése esetén  $N$  értéke 2 lesz, míg szavankénti beszéd esetében  $N$  értékét 5 és 10 közé szokás választani.
- Állapot-átmeneti valószínűségi mátrix  $\mathbf{A} = \{a_{ij}, i, j = 1, \dots, N\}$ . Ez biztosít kapcsolatot vagy Markov hálózatot hoz létre az egyes állapotok között, és modellezi a variációkat az állapotokhoz rendelt jelek időtartama alatt. A balról–jobbra rejtett Markov modell esetében  $a_{ij} = 0$ , minden  $i > j$  esetén, ezért ez az  $\mathbf{A}$  mátrix egy felső háromszög mátrix.
- Állapot tapasztalati vektorok  $\{\mu_{i1}, \mu_{i2}, \dots, \mu_{iM}, i = 1, \dots, N\}$ . Minden állapotra egy halmaz  $M$  prototípus vektor, a jel terének centroidjainak modellezésére minden állapotban.
- Állapot tapasztalati vektor valószínűségi modellje. Lehet diszkrét az  $M$  prototípusvektorokból, azok hozzárendelt sűrűségfüggvényével  $\mathbf{P} = \{P_{ij}(\cdot); i=1, \dots, N, j=1, \dots, M\}$ , vagy folytonos (általában normális)  $\mathbf{F} = \{f_{ij}(\cdot); i=1, \dots, N, j=1, \dots, M\}$  eloszlásfüggvény modell.
- Kezdőállapot valószínűségi vektor  $\boldsymbol{\pi} = [\pi_1, \pi_2, \dots, \pi_N]$

2.2.2.3. A rejtett Markov modell betanítása

Az első lépés, hogy adatbázist alakítsunk ki a modell számára, nagy számú minta vételezésével a vizsgálat tárgyát képező jelből. Tegyük fel, hogy az adatbázis  $L$  vektor értékekkel bíró szekvenciából áll  $[\mathbf{X}]=[\mathbf{X}_k; k=0, \dots, L-1]$ , és minden szekvenciában  $\mathbf{X}_k=[x(t); t=0, \dots, T_k-1]$  változó számú  $T_k$  vektor van. A paraméterek megfelelő beállításához a modell statisztikájának adathalmazát kell előállítani. Valószínűségi értelemben a modell alkalmassága a modell utólagos valószínűségével  $P_{M|X}(M|\mathbf{X})$  adható meg,  $M$  modell  $\mathbf{X}$  adathalmazára. Ez a betanító folyamat az  $M$  modell utólagos valószínűségének és az  $[\mathbf{X}]$  adathalmaznak a maximalizálását célozza, a Bayes szabály alkalmazásával:  $P_{M|X}(M|\mathbf{X}) = \frac{1}{f_X(\mathbf{X})} f_{X|M}(\mathbf{X}|M)P_M(M)$ ,

ahol a nevezőben álló  $f_X(\mathbf{X})$  egy normalizáló hatással bír, és  $P_M(M)$  az  $M$  modell korábbi valószínűsége. Az iménti egyenlet maximalizálása egyenértékű a valószínűségi függvény maximalizálásával, mely a következő egyenlet segítségével fejezhető ki  $\mathbf{X}$  vektorszekvenciára és  $M$  modellre:

$$f_{X|M}(\mathbf{X}|M) = \sum_s f_{X|S,M}(\mathbf{X}|s, M)P_{s|M}(s|M),$$

ahol  $f_{X|S,M}(\mathbf{X}(t)|s(t), M)$  jelöli az eloszlásfüggvényt az  $M$  modell  $\mathbf{X}$  szekvenciájának  $\mathbf{s}=[s(0), s(1), \dots, s(T-1)]$  állapotában, mely a következőképpen fejezhető ki:

$$f_{X|S,M}(\mathbf{X}|s, M) = f_{X|S}(\mathbf{x}(0)|s(0))f_{X|S}(\mathbf{x}(1)|s(1))\cdots f_{X|S}(\mathbf{x}(T-1)|s(T-1)),$$

ahol az  $N$  állapot egyike lehet  $s(t)$   $t$  időpillanatban,  $f_{X|S}(\mathbf{x}(t)|s(t))$  pedig

$f_{X|S,M}(\mathbf{X}(t)|s(t), M)$  rövidített leírása az  $M$  modell  $\mathbf{x}(t)$  eloszlásfüggvényének  $s(t)$

állapotban. A  $\mathbf{s}$  állapotszekvencia Markov valószínűsége a következőképpen adható

meg:  $P_{s|M}(s|M) = \pi_{s(0)}a_{s(0)s(1)}a_{s(1)s(2)}\cdots a_{s(T-2)s(T-1)}$ . Az előző három egyenletből az

utolsó kettőt behelyettesítve az elsőbe kapjuk a következőt:

$$\begin{aligned} f_{X|M}(\mathbf{X}|M) &= \sum_s f_{X|S,M}(\mathbf{X}|s, M)P_{s|M}(s|M) = \\ &= \sum_s \pi_{s(0)}f_{X|S}(\mathbf{x}(0)|s(0))a_{s(0)s(1)}f_{X|S}(\mathbf{x}(1)|s(1))\cdots a_{s(T-2)s(T-1)}f_{X|S}(\mathbf{x}(T-1)|s(T-1)) \end{aligned}$$

ahol az összegzés végighalad a összes  $\mathbf{s}$  állapotszekvenciát. A betanítási folyamatban az átmeneti valószínűségeket és a tapasztalati eloszlásfüggvények paramétereit az úgy kell kiszámítani, hogy iménti egyenletből a modell valószínűsége maximális legyen. A közvetlen maximalizálás a modell paramétereivel összefüggésben nem triviális megoldás. Következésképpen az egyenlet számításigénye egy megfigyelési

szekvencia  $T$  darab vektorára  $O(N^T)$  lesz. Ez még kis számok esetében is nagyon nagy számításigényt jelent. Az állapot-idő diagram ismétlődő struktúrái azt mutatják, hogy sok az ismétlődő számítás az egyenletben, melyet egy hatékony megvalósítással el tudunk kerülni.

### 3. Zajsűrési módszerek

#### 3.1. Zajzár

Érdekes, bár nem túl hatékony megoldás a jelszint folyamatos vizsgálata, és ha az elér egy kellően alacsony értéket, akkor a jelet teljesen kikapcsolva, csendet kapunk eredményül. Ezzel nem csak a zajt állítjuk meg, hanem az alacsony jelszintű hasznos jelet is, és a jel zajterhelése továbbra is megmarad, ha már egy adott szint felett van az értéke. Ez a módszer megoldást jelenthet olyan esetekben, mikor egy berendezés öngerjesztő folyamatának szeretnénk gátat szabni. Ilyen előfordulhat például mikrofonok esetén, melyek egy légtérben vannak az erősített jelüket megszólaltató hangsugárzókkal, és elhelyezéssel már nem tudunk a problémán segíteni (nagy hangerőre van szükség, és bárhová helyezzük a mikrofont, egy zárt térben, beindul az oszcilláció). Másik szemléletes példa az elektromos gitárok torzító berendezése, melynek alapvető működéséhez tartozik az átlagosnál nagyobb erősítés. Ilyenkor az érkező jel kikapcsolása a húrok öngerjesztő rezgésének megakadályozását szolgálhatja. Egy ilyen berendezés megfelelő beállítása nehéz, főleg, ha több vagy többféle hangforrás is van, ráadásul a valódi problémára nem jelent megoldást.

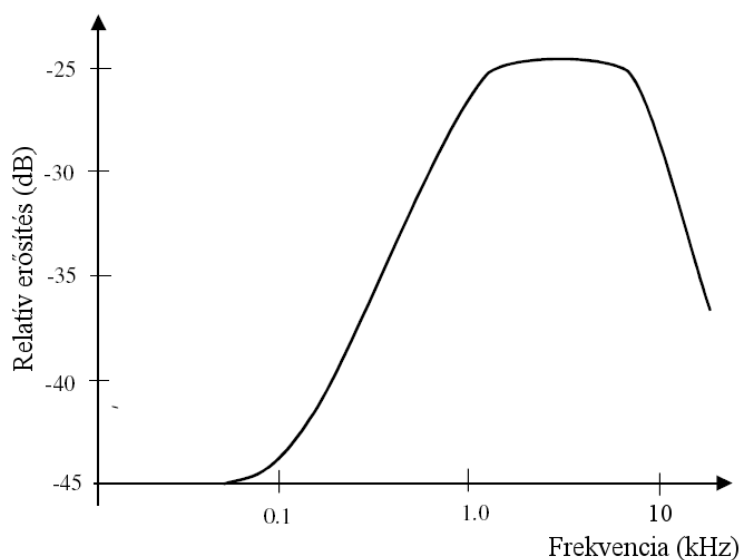
#### 3.2. Analóg valós idejű zajsűrők

##### 3.2.1. Dolby zajsűrő

A Dolby zajsűrő rendszerek úgy működnek, hogy emelik a magasabb frekvenciákon a hasznos jel amplitúdóját, ezzel javítva a jel–zaj arányt ezen frekvenciákon. A jel nagy része főként az alacsonyabb frekvenciákra – 2 kHz alá – koncentrálódik. A magasabb frekvenciájú jelek, melyek a minőséget és az élményt hordozzák viszonylag alacsonyabb energiájúak, és tovább degradálódnak a zajjal. Például egy mágneses szalagra rögzített jelsorozat esetében a szalag zaja és a korrekciós előerősítő fokozat zaja jelentős magasabb frekvenciájú „sistergést” állít elő. Lejátszásakor ezért ezeken a frekvenciákon romlik a jel–zaj arány – tehát nem lesz egyenletes a teljes átvitt sávban, vagyis az alacsonyabb frekvenciákon jobb, a magasabbakon rosszabb értéket kapunk. Tehát a magasabb frekvenciákon kevesebb hallható hasznos jelet kapunk. A Dolby zajsűrő tehát azon az alapelven működik, hogy a viszonylag kisebb energiájú magasabb frekvenciájú jeleket kiemeli, erősíti a hanganyag felvétele során, a többi jelet pedig változatlanul hagyja. Felvétel közben a jel

egy kiemelés és egy dinamika kompresszió keverékén megy keresztül, melyet lejátszáskor vissza kell alakítani az eredeti állapotába. A kódoló és a dekódoló rendszernek jól illesztettnek kell lennie, a torzítások elkerülése érdekében.

A Dolby laboratóriumban többféle ilyen szűrőt fejlesztettek ki, melyek a Dolby A, Dolby B, és Dolby C. Ezek alapjában véve a sávok számában és az alkalmazott kiemelő stratégiájában különböznek egymástól. A Dolby A a professzionális felhasználás számára lett kifejlesztve, és négy nagy részre osztja a frekvenciaspektrumot – az első sáv 0–80 Hz (alul áteresztő szűrő), a második sáv 80–3000 Hz (hullámsávszűrő), a harmadik sáv 3 kHz körüli felül áteresztő szűrő, a negyedik sáv pedig egy 9 kHz körüli felül áteresztő szűrő segítségével jött létre. A kódoló minden hullámsáv erősítését alkalmazkodóan igazította, az alacsony energiájú jelek kiemelése érdekében. A Dolby A rendszer gondoskodik a 10–15 dB-es erősítésről minden sávban, ha a jel 45 dB-nél nagyobb értékkel esik a maximális rögzítési jelszint alá. A kereskedelmi forgalomba kerülő készülékekben a Dolby B és a Dolby C zajsűrőket alkalmazták. Ezek a rendszerek az imént ismertetett Dolby A rendszerrel szemben már csak két hullámsávon dolgoznak, a négy helyett. A



17. ábra A Dolby C rendszer elő-kiemelő sávátvitele

Dolby B rendszer 10 dB-es emelést biztosít alacsony jelszint esetén (nagyobb, mint 45 dB eltérés a maximális rögzítési jelszinthez képest), a Dolby C rendszer pedig 20 dB-es emeléssel dolgozik, ahogy az a következő ábrán látható.

Ez a módszer csak analóg jelek esetében egy jól meghatározott zajforma csökkentésére alkalmas,

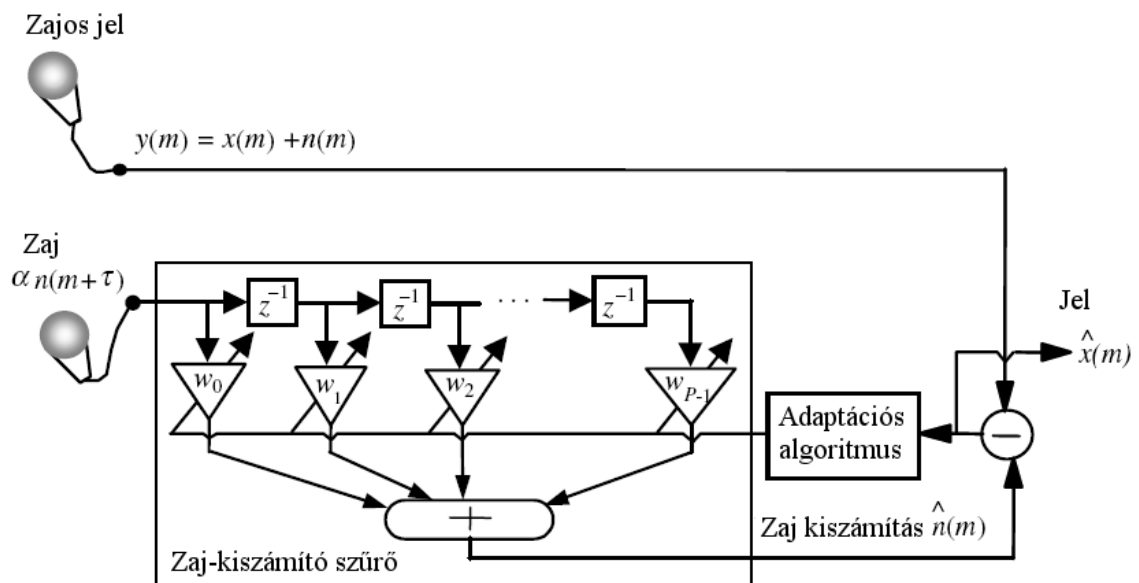
ezért ez nagyon behatárolja az alkalmazási területet. Ezért alkalmazzák ezt a módszert szalagos rögzítő eszközök – úgy mint kompakt kazettás magnetofonok – esetében. Az ilyen készülékek esetében jellemző, hogy a szalagon található mágneses részecskék rendezettsége és az előmágnesezés milyensége befolyásolja a hangminőséget, továbbá a magasabb frekvenciájú hangjelek – az alkalmazott mágne-

szehető anyag függvényében – kisebb jelszinten érkeznek vissza a szalagról, mely egy korrekciós előerősítő fokozat alkalmazását teszi szükségessé, ezáltal az elektronikus rendszernek is lesz egy zaja, melyek együttesének „szűrésére” alkalmazható a Dolby A, B és C rendszer.

### 3.2.2. Zajérvénytelenítés

Egy zajos környezetből (mint például egy mozgó autó) származó beszéd esetén a hasznos jelet a környezet zajával együtt észleljük. A mérőrendszerek esetében az információhordozó jel gyakran a berendezést körülvevő környezet zajával szennyezett. A zajjal együtt észlelt  $y(m)$  jel a következőképpen modellezhető:  $y(m) = x(m) + n(m)$ , ahol  $x(m)$  jelenti a hasznos jelet,  $n(m)$  jelenti a zajt,  $m$  pedig a diszkrét idő mérőszámot. Néhány esetben, például mikor mobiltelefont használunk egy mozgó autóban, vagy rádiókommunikációs eszközt egy repülőgép pilótafülkéjében, talán lehetséges, hogy kimérhető és kiszámítható a környezeti zaj pillanatnyi amplitúdója, egy irányított mikrofon használatával. Ezután a az  $x(m)$  hasznos jel visszaállítható lehet a zajos jel és a zaj különbségeként.

Az adaptív zaj-érvénytelenítő rendszer – két bemenettel rendelkezik, melyekből az egyik a környezeti zajt érzékeli, a másik pedig a zajjal terhelt hasznos jelet – alkalmazható beszédhang kiemelésre. Ebben a rendszerben az egyik irányított mikrofon a zajos  $x(m) + n(m)$  jelet érzékeli, míg a másik irányított mikrofon az előzőtől egy adott távolságra elhelyezve a zajt  $\alpha \cdot n(m + \tau)$  méri. Az a csillapítási szor-



18. ábra Adaptív zajérvénytelenítő elvi felépítése



zótényező és a  $\tau$  időkéleltetési érték meglehetősen túlegyszerűsített modellt ad a mikrofon elhelyezésétől eltérő helyről származó és eltérő zajok hatásainak modellezésére. A második mikrofonból érkező zajt egy adaptív digitális szűrővel dolgozzuk fel, annak érdekében, hogy amplitúdója megegyezzen az első mikrofonban érzékelt hang zajszennyezésének amplitúdójával, ezután pedig ezt a jelet kivonjuk zajos jelből, a zaj érvénytelenítése céljából. Ez a berendezés a zaj alacsonyabb frekvenciáin működik hatékonyan, de szenved a jelek nem állandó jellegétől, és a túlegyszerűsített feltevéstől mely szerint lineáris szűrővel modellezhető a zaj szóródása és terjedése a térben. A berendezés felépítését a 18. ábra szemlélteti.

Sok esetben, például egy telekommunikációs rendszerben alkalmazott vevőkészülék esetében nem ismerjük a zajszennyezés pillanatnyi értékét, csak a zajos jel adott. Ezen esetekben nem érvényteleníthető a zaj, csak mértéke csökkenthető, a jel és a zaj statisztikáját felhasználva. A Wiener szűrő alkalmas ilyen jellegű zajcsökkentési eljárásra, ahol csak a zajos jel adott.

### **3.3. Digitális jelek feldolgoása**

#### *3.3.1. Átalakítások*

##### 3.3.1.1. Analóg–digitális átalakítás

Az analóg jelfeldolgozás során valamilyen folytonosan változó fizikai mennyiséget alakítunk elektronikusan feldolgozható jellé. Ilyen átalakítás történik például egy mikrofonban, ahol a levegő rezgéseit, nyomásváltozásait alakítjuk elektromos feszültséggé, melyet a későbbiekben elektronikus jelerősítő berendezésekkel tudunk tovább feldolgozni. Ezek a jelek még elég sérülékenyek, könnyen terhelődhetnek zajjal, és ezen zajok eltávolítása az analóg jelekből nagyon nehéz feladat. Ha a jelet szeretnénk megszabadítani az additív zajoktól, akkor az első lépés, amit meg kell tennünk a jel átalakítása digitális jellé. A digitális jel fontos jellemzője, hogy egy előre meghatározott értékhalmból vehet fel csak értékeket. Ezen jelek esetében a feldolgozás már matematikai modellek alapján digitális számítógépek segítségével történhet. Továbbá nagyon előnyös tulajdonságaik közé tartozik, hogy sokkal kevésbé érzékenyek a külső additív zajokra, a diszkrét értékek miatt. A legegyszerűbb ilyen digitális jelek kétértékűek, vagy más néven bináris jelek. A mai korszerű kommunikációs rendszerekben gyakran alkalmaznak olyan készülékeket, melyek az analóg jelet először csak digitálissá alakítják, majd ez a digitális jelet

alakítják tovább egy bitfolyammá, melynek értelmezése a vevő oldalon történik. Ez a bit soros adatátvitel, napjaink egyik legkedveltebb átviteli formája.

Az analóg–digitális átalakítás vagy konverzió (Analog to Digital Conversion – ADC) úgy történik, hogy egy analóg feszültségjelet alapul véve abból adott időközönként mintát veszünk. Ez a digitális jel egyik fontos jellemzője, a *mintavételezési frekvencia*. A legjobb minőségű digitális jelet akkor kapnánk, ha egy időegység alatt végtelen sokszor vennénk mintát az analóg jelből, de ez technikailag kivitelezhetetlen és értelmetlen is. A mintavételezés a gyakorlatban rövid időközönként történik:

$T_s = \frac{1}{F_s}$ , ahol  $T_s$  az időközt,  $F_s$  pedig a frekvenciát jelöli. Ez a *mintavételezési eljárás*

az  $x(t)$  analóg jelet egy  $x(m)$  impulzussorozattá alakítja.

Az átalakítás következő lépése a kvantálás, amikor a mintavételezés után kapott impulzusokat egy diszkrét értékhalmozra képezzük. A halmaz elemeit egy analóg jeltartomány meghatározott számú részekre osztásával kapjuk meg. A kvantálás során a mintavételezés eredményeképpen kapott impulzussorozat eleméről eldöntjük, hogy melyik két érték közé (melyik tartományba) esik, és előállítjuk a  $x(m)$  jelből az  $y(m)$  jelet, mely a kvantálás eredményeképpen keletkezik. Ez a digitális jel másik fő jellemzője, amit *kvantálási hossz*nak mondunk. A jel pontossága és a visszaalakításkor kapott eredmény nagymértékben függ a tartományok számától, így a kvantálási hosszról.

A következő lépés a kódolási művelet. Ebben a műveletben minden mintavételi pontban felvett jelhez a kvantálás végeredményének megfelelő bináris kódszót rendelünk hozzá. Ezzel az eljárással minden  $y(m)$  jelhez bináris kódszót rendeltünk, tehát végeredményként kaptunk egy digitálisan feldolgozható jelet.

### 3.3.1.2. Digitális–analóg átalakítás

A digitális jel analóg jellé történő alakításakor (Digital to Analog Conversion – DAC) először a vett bináris jelsorozatot kell helyreállítani, amennyiben az szükséges. A jelsorozat visszaállítását követően a kódszavakból előállítható az  $y'(m)$  függvény, melyet egy alul-áteresztő szűrővel szűrve megkapható az eredeti  $x(t)$  jelet közelítő  $x'(t)$  analóg jel.

A közelítés javítására alkalmazhatunk interpolációt, mely a két függvény hasonlóságát növeli és csökkenti a torzítást. Az analóg jel mintavételezett formájából történő visszaállítás feltétele a Nyquist elméletben megadott. A Nyquist elmélet azt mondja ki, hogy egy sávkorlátozott jel esetében annak  $F_c$  legnagyobb frekvenci-

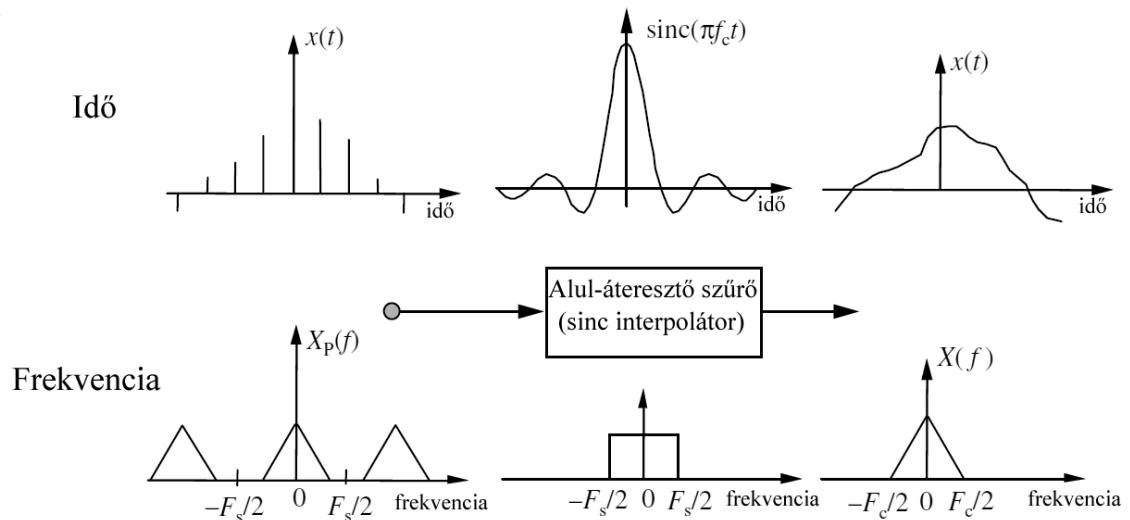
ájú összetevője is visszaállítható abban az esetben, ha a mintavételezés másodpercenkénti frekvenciája nagyobb volt a frekvencia kétszeresénél, vagyis  $F_s > 2F_c$ . Tegyük fel, hogy egy sávkorlátozott  $x(t)$  folytonos jel  $F_s$  másodpercenkénti mintavételezési frekvenciával mintavételezett. A diszkrét időközönként mintavételezett jel  $x(m)$  a következő szorzattal fejezhető ki:  $x(m) = x(t)p(t) = \sum_{m=-\infty}^{\infty} x(t)\delta(t - mT_s)$ , ahol

$p(t) = \delta(t - mT_s)$  a mintavételező függvény, és  $T_s = \frac{1}{F_s}$  a mintavételezési idő intervallum hossza. Vegyük az iménti egyenletet Fourier transzformálva, mely megmutatja, hogy a mintavételezett jel spektruma a következő egyenlettel írható fel:

$X_s(f) = X(f) * P(f) = \sum_{k=-\infty}^{\infty} X(f + kF_s)$ , ahol  $X(f)$  az  $x(t)$  jel spektruma, és  $P(f)$  a  $p(t)$

mintavételező függvénye, a  $*$  pedig a konvolúciós szorzás műveletét jelöli. Az iménti egyenlet állítása szerint a mintavételezett jel spektruma az alapsáv  $X(f)$  spektrumából és annak egyenletes  $T_s = \frac{1}{F_s}$  közönkénti ismétléséből vagy leképezéséből

áll össze. Ha a mintavételezési frekvencia felette van a Nyquist aránynak, akkor az alapsáv  $X(f)$  spektruma nincs átfedésben a  $X(f \pm kF_s)$  képeivel, és az eredeti jel



19. ábra Folytonos jel visszaállítása mintáiból. A frekvenciatartomány interpolációja az alul-áteresztő szűrővel ekvivalens

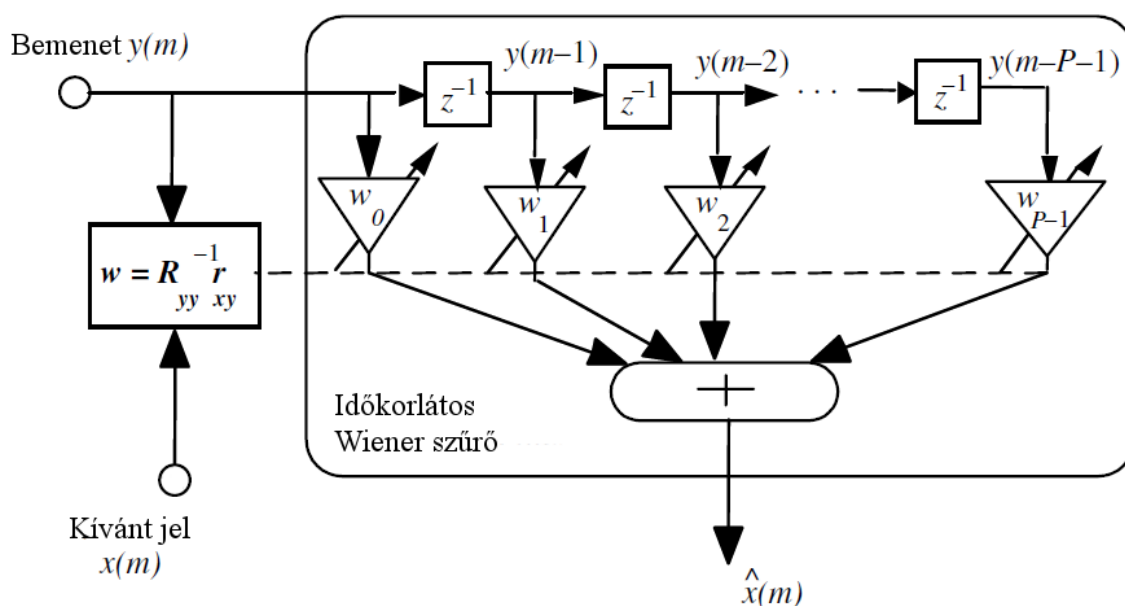
A gyakorlatban az  $F_s$  mintavételezési frekvencia nagyobb mint  $2F_s$ , általában  $2,5F_s$ , annak érdekében, hogy az átmeneti sáv szélességet hozzáigazítsuk az interpoláció alul-áteresztő szűrőjéhez.

### 3.3.2. A Wiener szűrő

A Wiener elméletet kidolgozójáról, Norbert Wiener-ről nevezték el, aki megfogalmazta az adatfüggő lineáris legkisebb négyzetek módszerére épülő szűrő elméletét. A Wiener szűrő központi szerepet játszik alkalmazásainak széles körében, mint például a visszhang érvénytelenítés, jelvisszaállítás, csatorna kompenzáció és rendszerazonosítás. A szűrő együtthatóit úgy kell kiszámítani, hogy a szűrő kimenő jele a lehető legjobban hasonlítson – vagyis különbségük négyzetének várható értéke legyen minimális – egy ismert úgynevezett megkívánt jelre. A Wiener elmélet feltételezi, hogy a jelek állandóak. Azonban ha a szűrő együtthatóit időnként újra kiszámítjuk, minden  $N$  elemű jelblokk után, a szűrő alakítja önmagát, a blokkokon belüli átlagos jellemzőkhöz, így blokk-alkalmazkodóvá válik. A blokk-alkalmazkodó (vagy szegmens-alkalmazkodó) szűrő alkalmazható olyan jelek esetében, mint például az emberi beszéd.

#### 3.3.2.1. A legkisebb négyzetek becslése

Wiener eredeti munkájában a folyamatos idejű legkisebb négyzetek módszerét adta meg, interpolációval. A módszert később kiterjesztette diszkrét időre is, mely egyszerűen megoldható, és a gyakorlatban jobban hasznosítható a digitális jelpro-



20. ábra Wiener szűrő felépítése

cesszorokban. A szűrő lehet végtelen hosszú jel feldolgozását célzó, vagy véges jelhosszúságot figyelembevevő. Általában a végtelen szűrőnek egy halmaz nem lineáris egyenlet az eredménye, a véges hosszúságú jelre alkalmazott szűrőnek pedig egy halmaz lineáris egyenlet, mely egy zárt alakú megoldás. A véges szűrő viszonylag egyszerűbb számítást tesz lehetővé, működéséből adódóan stabilabb és praktikusabb is. Hátránya a végtelen szűrővel szemben, hogy a kívánt eredmény elérése érdekében nagyon sok együttható szükséges.

A 20. ábrán szemléltetett Wiener szűrő egy  $\mathbf{w}$  együttható vektorral megvalósított. A szűrő kap egy  $y(m)$  bemeneti jelet, és készít egy  $\hat{x}(m)$  kimeneti jelet, mely a legkisebb négyzetes eltérést mutatja az  $x(m)$  megkívánt vagy céljelhez képest. A szűrő bemenete és kimenete közötti összefüggés a következő egyenlettel írható fel:

$$\hat{x}(m) = \sum_{k=0}^{P-1} w_k y(m-k) = \mathbf{w}^T \mathbf{y}, \text{ ahol } m \text{ a diszkrét idő mérőszám, a szűrő bemeneti}$$

jele:  $\mathbf{y}^T = [y(m), y(m-1), \dots, y(m-P+1)]$ , és a paramétervektor (a Wiener szűrő együttható vektora)  $\mathbf{w}^T = [w_0, w_1, \dots, w_{P-1}]$ . A szűrési művelet kétféle egymással ekvivalens alakban fejthető ki, konvolúciós összegként, és a belső vektorok szorzataként. A hibajel  $e(m)$  a kimenőjel  $\hat{x}(m)$  és a megkívánt jel  $x(m)$  különbségeként számítható ki a következő összefüggéssel:  $e(m) = x(m) - \hat{x}(m) = x(m) - \mathbf{w}^T \mathbf{y}$ . Ebből az egyenletből kiderül, hogy egy adott  $y(m)$  bemenőjelre és  $x(m)$  megkívánt jelre a hibajel a  $\mathbf{w}$  együtthatóvektortól függ. A hibajel és az együtthatóvektor közötti összefüggés kiderítése érdekében kifejtve az előző egyenletet a következőt kapjuk:

$$\begin{pmatrix} e(0) \\ e(1) \\ e(2) \\ \vdots \\ e(N-1) \end{pmatrix} = \begin{pmatrix} x(0) \\ x(1) \\ x(2) \\ \vdots \\ x(N-1) \end{pmatrix} - \begin{pmatrix} y(0) & y(-1) & y(-2) & \dots & y(1-P) \\ y(1) & y(0) & y(-1) & \dots & y(2-P) \\ y(2) & y(1) & y(0) & \dots & y(3-P) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y(N-1) & y(N-2) & y(N-3) & \dots & y(N-P) \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_{P-1} \end{pmatrix}.$$

Ez az egyenlet  $N$  mintára,  $y(m)$  és  $x(m)$  jelekre lett felírva. Tömörített alakban a következőképpen írható fel:  $\mathbf{e} = \mathbf{x} - \mathbf{Y}\mathbf{w}$ , ahol  $\mathbf{e}$  a hibavektor,  $\mathbf{x}$  a kívánt jel vektor,  $\mathbf{Y}$  a bementi jel mátrix, és  $\mathbf{Y}\mathbf{w}$  szorzat pedig a  $\hat{x}$  kimenőjel vektor. Feltételezzük, hogy  $P$  kiinduló minta ismert, vagy hanem akkor 0 értékűnek tekintendő. Ha az előbb felírt mátrixegyenletben az  $N$  és a  $P$  értéke egyenlő, akkor egy négyzetes mátrixegyenletet kapunk, és ez az egyetlen megoldás, ahol a számítási hiba értéke nulla. Ha  $N < P$  a minták száma  $N$  elégtelen mennyiségű ahhoz, hogy elérjük az egyedülálló megoldást a szűrő együtthatóinál, ezért ebben az esetben végtelen sok megoldás létezik nulla számítási hibával és a mátrixegyenletet határozatlannak mond-

juk. A gyakorlatban a jelminták száma sokkal nagyobb, mint a szűrő hossza – vagyis  $N > P$  –, ebben az esetben a mátrixegyenletet túlhatározottnak mondjuk, és létezik egy egyedi megoldás, általában nulla számítási hibával. Mikor  $N > P$  a szűrő együtthatói úgy lesznek kiszámítva, hogy az átlagos hiba minimális legyen, mint például a hiba átlagos értékének abszolút értéke  $E[|e(m)|]$ , vagy a hiba négyzetének középértéke  $E[e^2(m)]$ , ahol  $E[\cdot]$  a várható érték operátor. A hibafüggvény kiválasztása hatással van a megoldás optimális mivoltára, és a számítási komplexitására. A Wiener elméletben az objektív kritérium a legkisebb négyzetes hiba (LSE<sup>2</sup>) a szűrő kimeneti jele és a kívánt jel között. A legkisebb négyzetes hiba kritérium optimális a normális eloszlást követő jelek esetében. A korlátozott idő intervallumra alkalmazott legkisebb négyzetes hiba kritériumú Wiener szűrő egy lineáris, zárt formulához vezet. A Wiener szűrő együtthatói egy átlagos négyzetes hiba függvény  $E[e^2(m)]$  minimalizálásával eszközli ki, tekintettel a szűrő  $\mathbf{w}$  együttható vektorára. A számítási hiba négyzetének középértéke a következő egyenlettel adható meg:  $E[e^2(m)] = E[(x(m) - \mathbf{w}^T \mathbf{y})^2] = E[x^2(m)] - 2\mathbf{w}^T E[\mathbf{y}x(m)] + \mathbf{w}^T E[\mathbf{y}\mathbf{y}^T] \mathbf{w} = r_{xx}(0) - 2\mathbf{w}^T \mathbf{r}_{yx} + \mathbf{w}^T \mathbf{R}_{yy} \mathbf{w}$ , ahol  $\mathbf{R}_{yy} = E[\mathbf{y}(m)\mathbf{y}^T(m)]$  a bemenő jel autokorrelációs<sup>3</sup> mátrixa,  $\mathbf{r}_{xy} = E[x(m)\mathbf{y}(m)]$  pedig a bemeneti jel és a kívánt jel kereszt-kölcsönösségi vektora. Az előző kifejezést kifejtve a következő egyenletet kapjuk:

$$E[e^2(m)] = r_{xx}(0) - 2 \sum_{k=0}^{P-1} w_k r_{yx}(k) + \sum_{k=0}^{P-1} w_k \sum_{j=0}^{P-1} w_j r_{yy}(k-j), \text{ ahol } r_{yy}(k) \text{ és } r_{yx}(k) \text{ az } \mathbf{R}_{yy}$$

autokorrelációs mátrix és az  $\mathbf{r}_{xy}$  kereszt-kölcsönösségi vektor elemei külön-külön. Az előző egyenletből a korlátozott idő intervallumra alkalmazott szűrő négyzetes hibájának középérték számítására egy másodfokú függvény alkalmazható a  $\mathbf{w}$  együttható vektor értékeire, mely függvénynek egyetlen minimuma van. Például ha szűrőnek csak két együtthatója van ( $w_1, w_2$ ), akkor a négyzetes hiba középérték függvénye egy medence alakú képet mutat, egyetlen minimummal. A legkisebb négyzetes hiba jelenti az energiaminimumot.

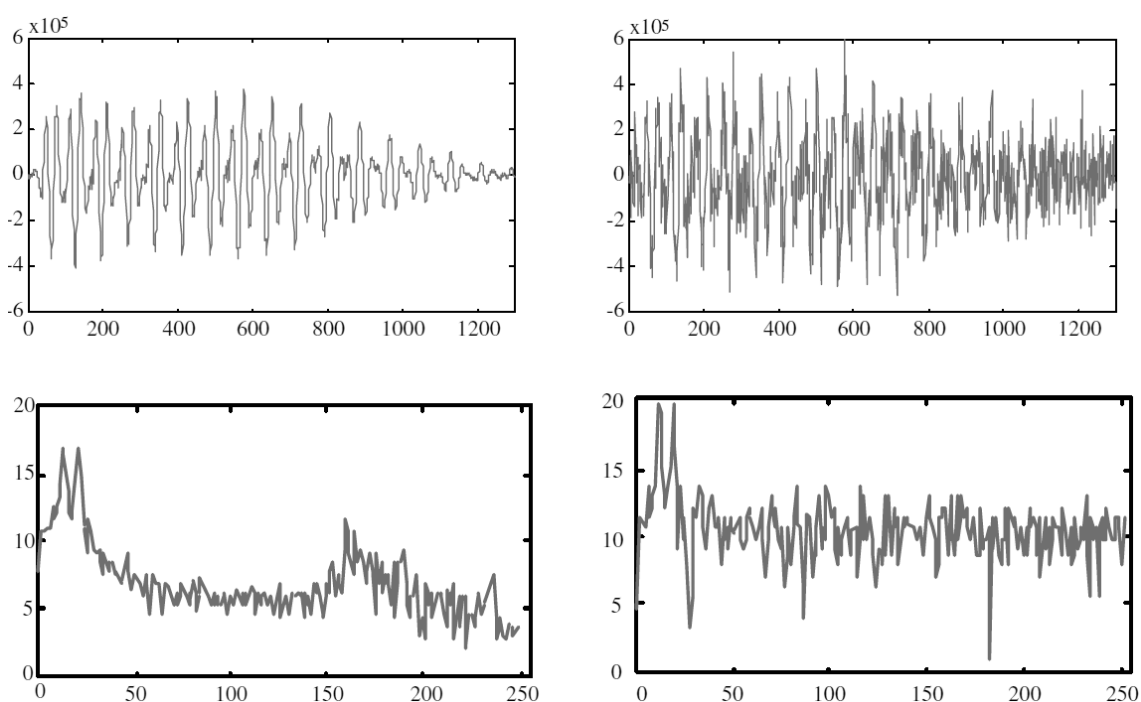
### 3.3.3. Zajcsökkentés spektrális kivonással

A spektrális kivonás módszerének lényege, hogy a zajos jelből veszünk egy olyan szakaszt, ahol a hasznos jel nincs jelen, tehát csak zaj észlelhető, és ezt a jelet egy megfelelő módszerrel kivonjuk a zajos jelből. Ez megtehető oly módon, hogy a jelek statisztikáit és spektrális paramétereit vesszük alapul. Ezzel a módszerrel adott esetben jelvisszaállítás is lehetséges.

A problémára a legegyszerűbb megoldást jelentené, ha külön rendelkezésre állna a zajos jel és a zaj. Ebben az esetben egy egyszerű művelettel eltávolítható lenne a zaj. A zajérvénytelenítő, melynek működését már korábban kifejtettem pontosan ilyen elven működik. Általában azonban csak a zajos jel áll rendelkezésre, tehát nincs lehetőség a zaj érvénytelenítésére, de előfordulhat, hogy a zaj hatását csökkenteni tudjuk a hasznos jel spektrumán.

### 3.3.3.1. A spektrális kivonás módszere

A 21. ábra olyan jeleket mutat, melyek zajjal terhelték, de nincs lehetőség a zaj érvénytelenítésére. A hasznos jel spektrumának középvértéke növelhető úgy, hogy a

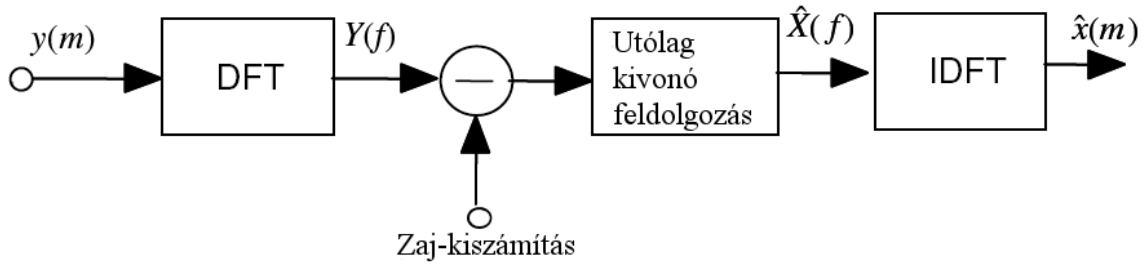


21. ábra Zajjal terhelt jelek alakjai az idő függvényében

zajos jel spektrumának középvértékéből kivonjuk a zaj spektrumának középvértékét. A zajos jel időben a következőképpen modellezhető:  $y(m)=x(m)+n(m)$ , ahol  $y(m)$  jelöli a zajjal terhelt jelet,  $x(m)$  a hasznos jelet,  $n(m)$  a zajt,  $m$  pedig a diszkrét idő mérőszám. Egy frekvenciatartományban az előző egyenlet a következőképpen fejthető ki:  $Y(f)=X(f)+N(f)$ , ahol  $Y(f)$  az  $y(m)$  zajos jel Fourier transzformáltja,  $X(f)$  az  $x(m)$  hasznos jel Fourier transzformáltja,  $N(f)$  az  $n(m)$  zaj Fourier transzformáltja, és  $f$  a frekvenciaváltozó. A spektrális kivonásban az  $y(m)$  zajos jelet vesszük, ezt egy átmeneti tárolóba helyezzük, majd felosztjuk  $N$  egyenlő részre. Minden szegmens esetében Hann vagy Hamming ablak segítségével szűrjük a spektrumot, ezután transzformáljuk diszkrét Fourier transzformációval (DFT) az összes  $N$  darab spektrális mintát. Az ablakok minden szegmens végén csökkentik annak hatását,

hogy a vett jelet feldaraboltuk, vagyis nem folytonos jeleket illesztünk össze. Az ablakozott jel így adható meg:  $y_w(m) = w(m)y(m) = w(m)[x(m) + n(m)] = x_w(m) + n_w(m)$ . Az alakozási eljárás a frekvenciatartományokra az alábbi módon fejthető ki (DFT segítségével):  $Y_w(f) = W(f) * Y(f) = X_w(f) + N_w(f)$ , ahol a  $*$  operátor a konvolúciós szorzást jelenti.

A következő ábra a spektrális kivonás módszerének blokkdiagramját szemlélteti. (Az IDFT jelölés a diszkrét Fourier transzformáció inverzét jelenti)



22. ábra A spektrális kivonás egyszerűsített blokkdiagramja

A spektrális kivonás egyenlettel a következő módon írható le:  $|\hat{X}(f)|^b = |Y(f)|^b - \alpha \overline{|N(f)|^b}$ , ahol  $|\hat{X}(f)|^b$  az eredeti jel spektrumának kiszámítása,  $|X(f)|^b$  és  $|N(f)|^b$  pedig az időátlagolt zaj spektrum. A spektrális kivonás nagyságához a  $b$  kitevő értékét 1-re, az erősségéhez pedig 2-re kell választani. Az  $\alpha$  paraméter a kivont zaj mennyiségét szabályozza. A teljes kivonáshoz  $\alpha=1$  értéket kell választani, a túlvonáshoz pedig  $\alpha>1$  érték lesz a megfelelő. Az időátlagolt zajspektrum beszerezhető a jelmentes szakaszokból, ahol csak kizárólag a zaj van jelen, nincs hasznos jel, és az alábbi összefüggés segítségével formalizálható:

$$\overline{|N(f)|^b} = \frac{1}{K} \sum_{i=0}^{K-1} |N_i(f)|^b, \text{ ahol } |N_i(f)|^b \text{ a spektrum } i\text{-edik zajkerete, és feltesszük,}$$

hogy a zaj  $K$  keretből áll, és  $K$  változó. Más esetben az átlagolt zajspektrum beszerezhető az elsőrendű digitális alul-áteresztő szűrő kimeneteként a következőképpen:  $\overline{|N_i(f)|^b} = \rho \overline{|N_{i-1}(f)|^b} + (1-\rho)|N_i(f)|^b$ , ahol az alul-áteresztő szűrő együtthatója  $\rho$ , melynek tipikus értéke 0,85 és 0,99 közé szokott esni. Az időtartomány alapú jelvisszaállításhoz a nagyságspektrum  $|\hat{X}(f)|$  kiszámítását kombináljuk a zajos jel fázisával, majd ezt követően transzformáljuk időtartományba az inverz diszkrét Fourier transzformáció segítségével, mely a következőképpen írható fel:

$$\hat{x}(m) = \sum_{k=0}^{N-1} |\hat{X}(k)| e^{j\theta_Y(k)} e^{-j\frac{2\pi}{N}km}, \text{ ahol } \theta_Y(k) \text{ a zajos jel fázisa } Y(k) \text{ frekvencián. Ez a jel-}$$



viSSzaállító egyenlet feltételezi, hogy a hallható tartományba eső zaj főként a nagyságspektrum torzítása és a fázistorzítás miatt nagyrészt nem hallható. A szimulált fázistorzítás megfigyelhető hatásának becslése igazolta a feltevést. A zaj spektrumának változatossága következtében előfordulhat, hogy nagyság és az energiaspektrum szempontjából negatív eredményt kapunk a spektrális kivonással. Ennek végeredményeképpen nagyon valószínű, hogy csökkenni fog a jel-zaj arány. A negatív eredmények elkerülése érdekében a spektrális kivonás kimeneti jelét egy utófeldolgozó térképező függvénnyel feldolgozzuk, melyet a következőképpen formalizálhatunk:  $T[|\hat{X}(f)|] = \begin{cases} |\hat{X}(f)| & \text{ha } |\hat{X}(f)| > \beta |Y(f)| \\ \text{fn}[|Y(f)|] & \text{egyébként} \end{cases}$ . Például választhatunk egy

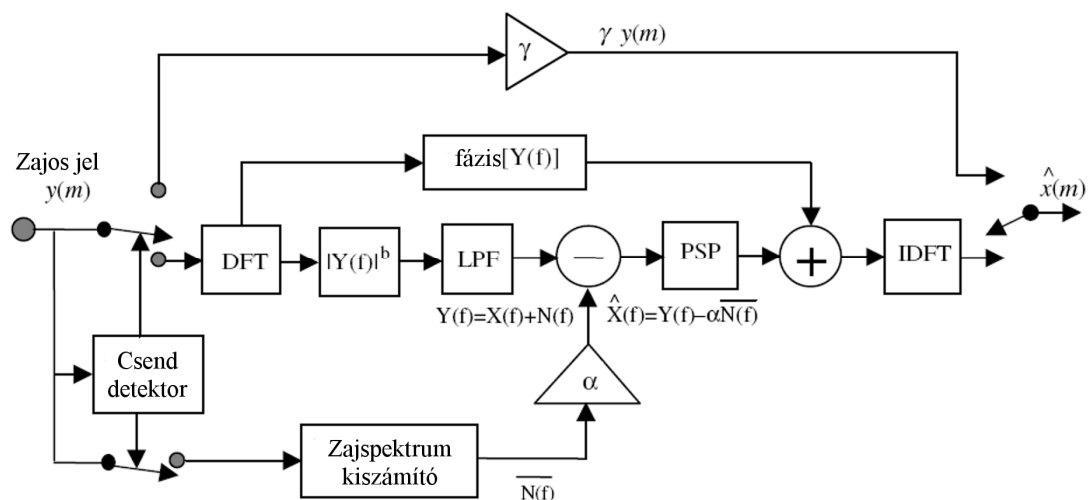
olyan szabályt, mely ha a  $|\hat{X}(f)| > 0,01 |Y(f)|$  (a nagyság spektrum szorzása 0,01-dal ekvivalens a  $-40\text{dB}$ -lel) érték igaz, akkor az  $|\hat{X}(f)|$  érték valamilyen függvénye legyen a zajos jelnek  $\text{fn}[|Y(f)|]$ . Ennek legegyszerűbb formája, hogy beállítunk egy  $\text{fn}[|Y(f)|] = \text{zajszint}$  értéket, ahol a zajszint pozitív konstans. Egy másik lehetőség, ha  $\text{fn}[|Y(f)|] = \beta |Y(f)|$ . Ez esetben:

$$T[|\hat{X}(f)|] = \begin{cases} |\hat{X}(f)| & \text{ha } |\hat{X}(f)| > \beta |Y(f)| \\ \beta |Y(f)| & \text{egyébként} \end{cases}$$

A spektrális kivonás energiaspektrum vagy nagyságspektrum tartományok alapján is implementálható. A két módszer hasonló, de elméletileg eredményeznek némi különbséget a várható teljesítményben.

### 3.3.3.2. A spektrális kivonó megvalósítása

Egy spektrális zajkivonó rendszer blokkdiagramját mutatja az alábbi ábra,



23. ábra Spektrális kivonó megvalósításának blokkdiagramja

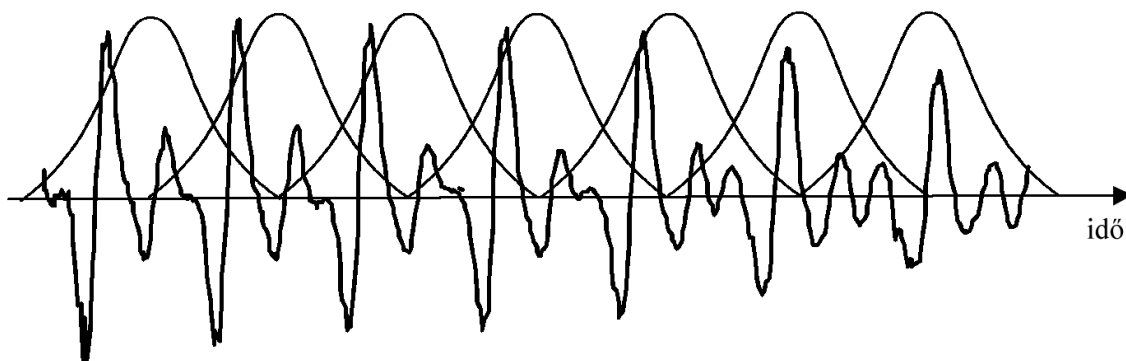
mely a következő alrendszeréből áll:

- a.) egy csend detektor a jelmentes időszakok megtalálására, a zaj spektruma ezen időtartamok alatt frissül.
- b.) a diszkrét Fourier transzformációt megvalósító rész (DFT) az időtartományt frekvenciatartományá alakítja, a DFT egységet egy nagyság operátor követi.
- c.) alul-áteresztő szűrő (Lowpass Filter, LPF), a zaj egyenetlenségeinek csökkentésére, melynek célja a zaj ingadozásaiból adódó feldolgozási torzítás csökkentése.
- d.) utófeldolgozó, a spektrális kivonás eredményeképpen keletkezett torzítások eltávolítására.
- e.) inverz diszkrét Fourier transzformációt végző rész, mely frekvenciatartományból viszi a feldolgozott jelet időtartományba.
- f.) csillapító berendezés  $\gamma$ , mely a zaj csillapítását végzi a csöndes részeken.

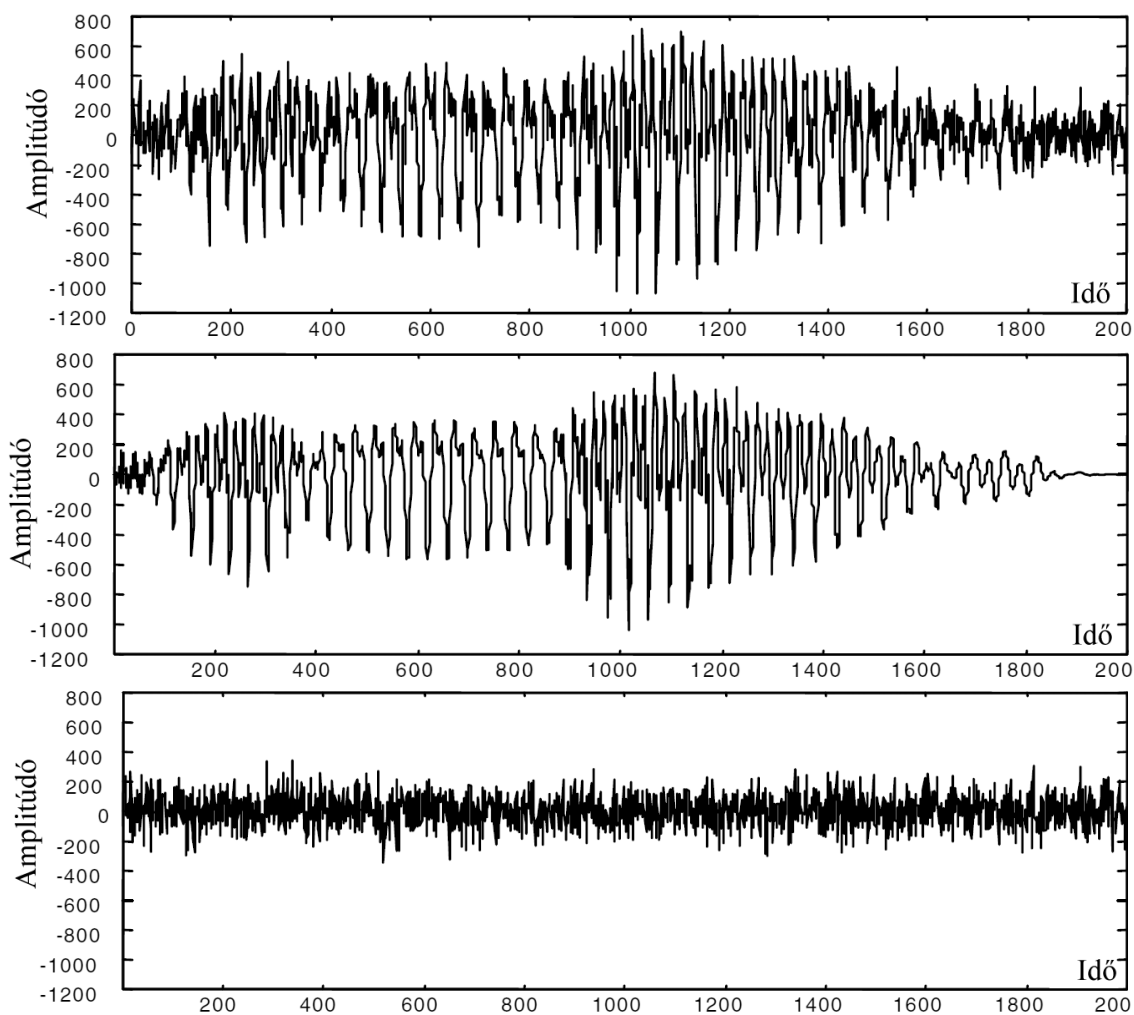
A diszkrét Fourier transzformáción alapuló spektrális kivonó blokkdiagramja egy algoritmus. A beérkező audiojelet átmeneti tárolóba rakjuk, majd  $N$  darab egymást átfedő mintablokkra osztjuk. Minden blokk Hann vagy Hamming ablakolt, és azt követően diszkrét Fourier transzformációval átalakítjuk frekvenciatartományá. A spektrális kivonás után a jel nagyságspektrumát kombináljuk a zajos jel fázisával, és visszaalakítjuk időtartományá. Minden blokkot egymáshoz illesztünk, úgy, hogy legyen átfedés a megelőző és a rákövetkező blokkal, és ez adja a végleges kimeneti jelalakot.

A blokk hosszának megválasztása a spektrális elemzéshez egy kompromiszsum, mely két ellentmondó szükséglet, az időfelbontás és a spektrális felbontás igényeit kell, hogy kielégítse. Tipikusan 5–50 ms közé esik a használt blokkok hossza. Ez egy 20 kHz-es mintavételezési frekvencián azt jelenti, hogy  $N$  értéke 100–1000 tartományba esik. A spektrum frekvenciafelbontása egyenesen arányos a minták számával ( $N$ ). Nagyobb  $N$  esetén a spektrum értékelése jobb lesz. Ez csak részben igaz a frekvenciatartomány alsó részére, mivel az alacsonyabb frekvenciájú részek az idő múlásával lassan változnak, és nagyobb ablak szükséges a stabil értékeléshez. Az ellentmondó szükséglet az audiojelek nem stacionárius természetéből adódnak, tehát az ablak méretének azért nem szabad túl nagyoknak lennie, hogy a rövid időtartamú események ne mosódjanak el.

Az ablakok átfedéseinek (24. ábra) fő feladata, hogy csillapítsák a blokkok végpontjain keletkező folytonossághiányokat, melyek a darabolásból adódnak. Bár

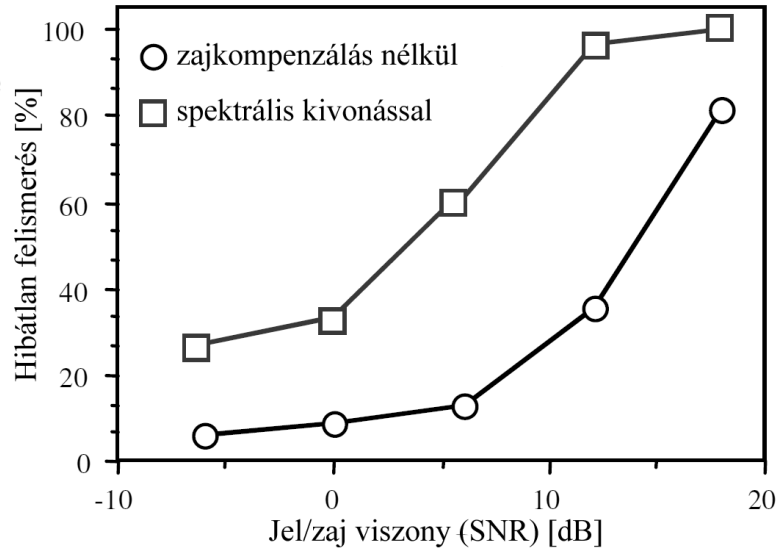


24. ábra A spektrális kivonásban alkalmazott ablakolási és átfedési folyamat létezik számos felhasználható ablaktípus, különböző karakterisztikákkal, mégis a legtöbb megvalósításban Hamming ablakokat használnak. A torzítás eltávolítás során – melyet a spektrális kivonás eredményezett – az utófeldolgozó algoritmus olyan információkat használ fel, mint a frekvenciacsatornák összefüggése az egymást követő blokkokban, valamint az jel eseményeinek és torzításának időtartamai.



25. ábra Zajos jel (felső), zajmentes jel (középső), és zaj (alsó)

A jel spektrális összetevőinek összefüggése – az időtengely mentén – részben szabályozható az ablakhossz megválasztásával és az átfedéssel. Ez az összefüggés növekszik az ablak hosszának csökkentésével és az átfedés növelésével. Az átfedés növelése azonban növeli a zajfrekvenciák összefüggését is, az időtengely mentén.



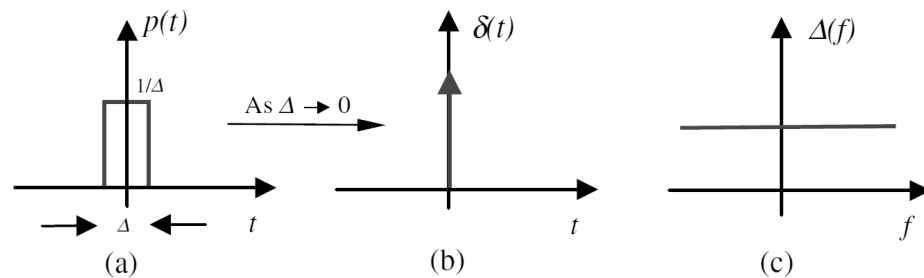
26. ábra A spektrális kivonás jelentősége a beszéd felismerésben

### 3.3.4. Az impulzív zaj eltávolítása

Az impulzív zajok gyors be/ki kapcsolásokhoz hasonló impulzusokat produkálnak és nagyon kiszámíthatatlan a megjelenésük, tehát bármely időpillanatban előfordulhatnak, és bármekkora időközrel. Ezen zajok eltávolítására és csökkentésére az előző részben kifejtett spektrális kivonás módszere nem alkalmas, mert itt a zajszint nem határozható meg.

#### 3.3.4.1. Impulzív zaj matematikai modellje

Az analóg impulzív zajt matematikai értelemben az alábbi ábra segítségével szemléltethetjük. Ha megvizsgáljuk az ábra (a) részén szemléltetett egységnyi te-



27. ábra (a) Egységnyi területű lökés, (b) impulzus, (c) az impulzus spektruma

rületű lökeshullám  $p(t)$  (négyzög jel) viselkedését, ahogy a lökés szélessége  $\Delta$  nullához tart, a lökés impulzussá válik. Az impulzus függvényt az ábra (b) része mutatja, és egy végtelenül kicsi időszélességgel rendelkező lökést a következőképpen

definiálhatunk:  $\delta(t) = \lim_{\Delta \rightarrow 0} p(t) = \begin{cases} \frac{1}{\Delta}, & |t| \leq \frac{\Delta}{2} \\ 0, & |t| > \frac{\Delta}{2} \end{cases}$ . Az impulzusfüggvény integrálja a kö-

vetkező:  $\int_{-\infty}^{\infty} \delta(t) dt = \Delta \times \frac{1}{\Delta} = 1$ . Az impulzusfüggvény Fourier transzformáltja felállítha-

tó:  $\Delta(f) = \int \delta(t) e^{-j2\pi ft} dt = e^0 = 1$ , ahol  $f$  a frekvenciaváltozó. Az impulzusfüggvényt

teszt jelleggel is szokták alkalmazni, a rendszer impulzusreakciójának megállapítására. Az ábra (c) részén mutatottak szerint az impulzus egy spektrálisan nagyon gazdag jel, minden frekvenciát azonos mennyiségben tartalmaz, tehát ezért alkalmas a tesztelésre.

Egy digitális impulzus  $\delta(m)$  (8. oldal 4. ábra (a) részén látható) úgy definiálható, mint egy egy minta hosszúságú „be” jel, és kifejezhető:  $\delta(m) = \begin{cases} 1, & m = 0 \\ 0, & m \neq 0 \end{cases}$ , ahol az

$m$  változó jelöli a diszkrét idő mérőszámot. A Fourier transzformációt használva

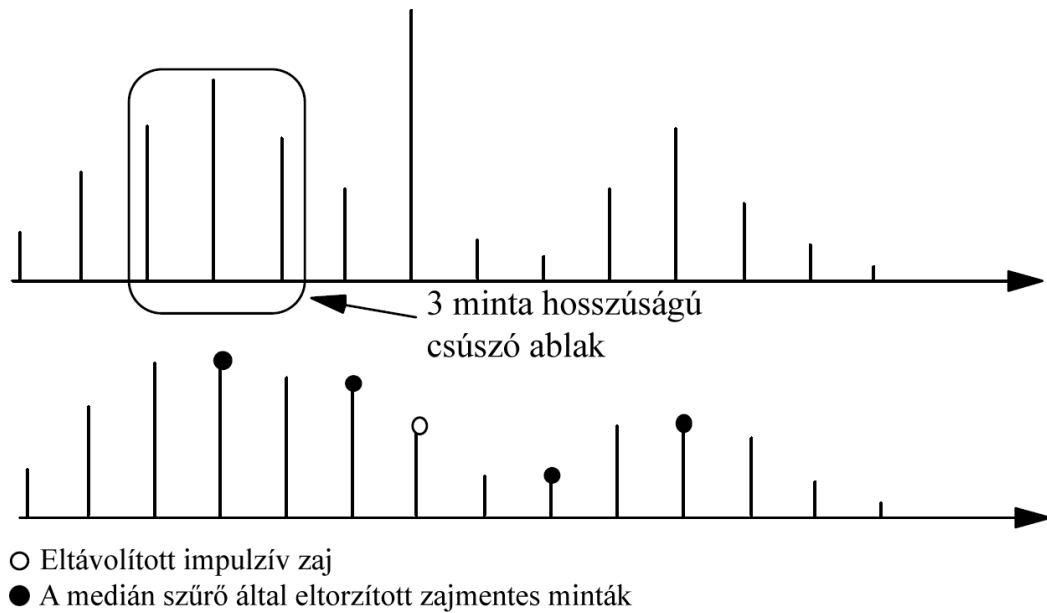
egy digitális impulzus frekvenciaspektruma megadható:  $\Delta(f) = \sum_{m=-\infty}^{\infty} \delta(m) e^{-j2\pi fm} = 1, 0,$

$-\infty < f < \infty$ .

Kommunikációs rendszerekbe a valódi impulzív típusú zajok hossza normálisan nagyobb, mint egy minta.

#### 3.3.4.2. A medián szűrő

Az impulzív zaj eltávolításának klasszikus megközelítése a medián szűrő. Medián alatt értjük az  $\{x(m)\}$  mintahalmaz elemeiből azt az  $x_{\text{med}}(m)$  értéket, melyre igaz, hogy a halmaz elemeinek fele ezen  $x_{\text{med}}(m)$  értéknél nagyobb, a másik fele pedig az  $x_{\text{med}}(m)$  értéknél kisebb. Tekintsük a mediánt minták egy halmazának, melyet úgy kapunk meg, hogy a mintákat rendezzük növekvő vagy csökkenő sorrendbe és kiválasztjuk a középső értéket. A medián szűrés esetén egy előre meghatározott hosszúságú ablakot csúsztatunk folyamatosan a jelen, és az ablakban található középső mintát kicseréljük az ablak mintáinak mediánjára, ahogy azt a következő ábra mutatja.



28. ábra Medián szűrő be- és kimenete. A működés közben torzítás is keletkezhet

A medián szűrő  $\hat{x}(m)$  kimenete az  $y(m)$  bemenet és egy  $2K+1$  mintát tartalmazó hosszúságú medián ablak alapján:

$$\hat{x}(m) = y_{\text{med}}(m) = \text{medián}[y(m-K), \dots, y(m), \dots, y(m+K)]$$

Egy számhalmaz mediánja a halmaz nem lineáris statisztikája egy nagyon jól hasznosítható tulajdonsággal, hogy nem érzékeny egy szokatlanul nagy mintaérték jelenlétére a halmazban, melyet kívülállónak hívunk. Ezzel szemben a középérték és részben a variancia érzékeny egy számhalmazban jelen levő kiugró értékre, vagyis az impulzív típusú zajra. Fontos tulajdonsága a medián szűrőnek – mely részben használható a képfeldolgozásban –, hogy megőrzi az éleket vagy a lépésenkénti folytonossághiányokat a jelben. A medián szűrők hatékonyan használhatók a képfeldolgozásban az impulzív zaj eltüntetésére az élek elmosása nélkül – ez nagyon jelentős a képfeldolgozásban. A medián filterrel történő kísérletezés során kiderült, hogy nem alkalmas magas minőségű eredmények produkálására audiojelek visszaállítása esetében. A medián szűrő nem tud mit kezdeni az olyan impulzív jellegű zajjal, mely hosszabb, mint egy vagy két minta. Továbbá a medián szűrő a hibásan impulzív zajnak vélt jelek módosításával jelentős mennyiségű feldolgozási torzítást okoz. A medián szűrő teljesítménye javítható egy alkalmazkodó határérték bevezetésével, tehát a mintát csak akkor cseréljük ki, ha annak eltérése a középső érték-

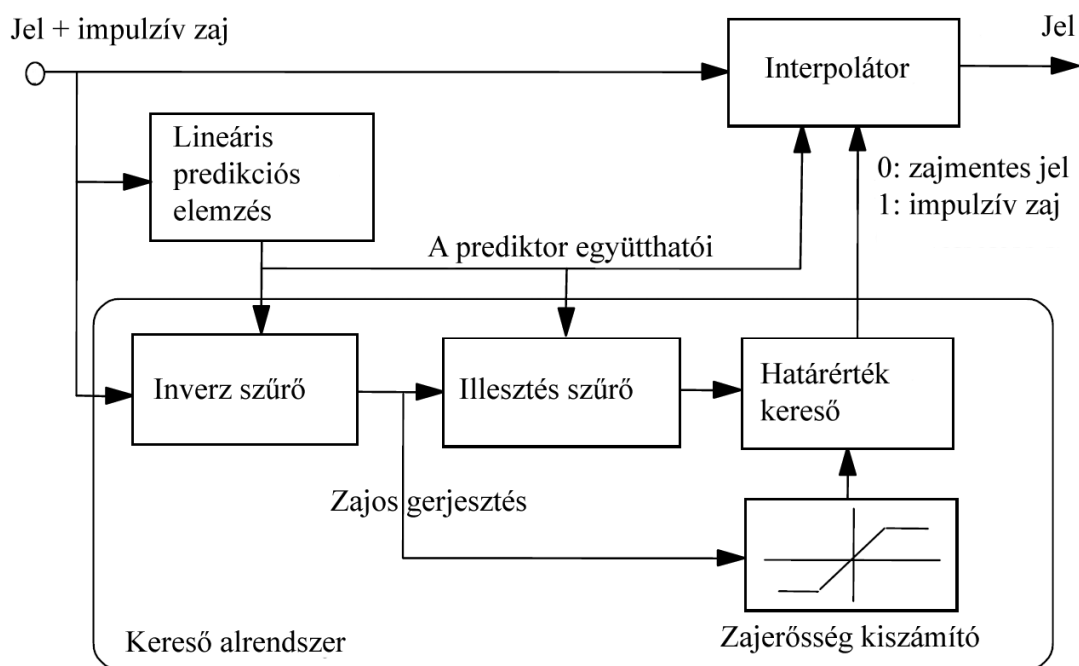
től meghaladja a határértéket:  $\hat{x}(m) = \begin{cases} y(m) & \text{ha } |y(m) - y_{\text{med}}(m)| < k\theta(m) \\ y_{\text{med}}(m) & \text{egyébként} \end{cases}$ , ahol

$\theta(m)$  egy alkalmazkodó határérték, mely erőteljes becsléssel határozza meg

$|y(m) - y_{\text{med}}(m)|$  átlagértékét,  $k$  pedig egy hangoló paraméter. A medián szűrők nem optimálisak, mert nem használhatók hatékonyan a jelgenerálás fiziológiájának előismerete vagy a jel és a zaj statisztikai eloszlásának modellezése nélkül.

### 3.3.4.3. Impulzív zaj eltávolítása lineáris prediktív modell segítségével

Az impulzív zavarok általában a jel egészét vizsgálva csupán csekély  $\alpha$  hányadát szennyezik. Mivel a jel jelentős  $1-\alpha$  része impulzív zajjal nem szennyezett, ezért előnyösebb megkeresni a zajos részeket (egyedülálló impulzusokat), és csak a torzult mintákat javítani. Ez a stratégia kiküszöböli a felesleges feldolgozásokat, és a minták egy viszonylag nagy része – mely nem szennyezett impulzív zajjal – érintetlenül eredeti állapotában marad. A következő ábra egy ilyen rendszer blokkdiagramját mutatja, mely két alrendszerből áll, egy keresőből és egy interpolátorból. A



29. ábra Egy impulzív zajeltávolító blokkdiagramja.

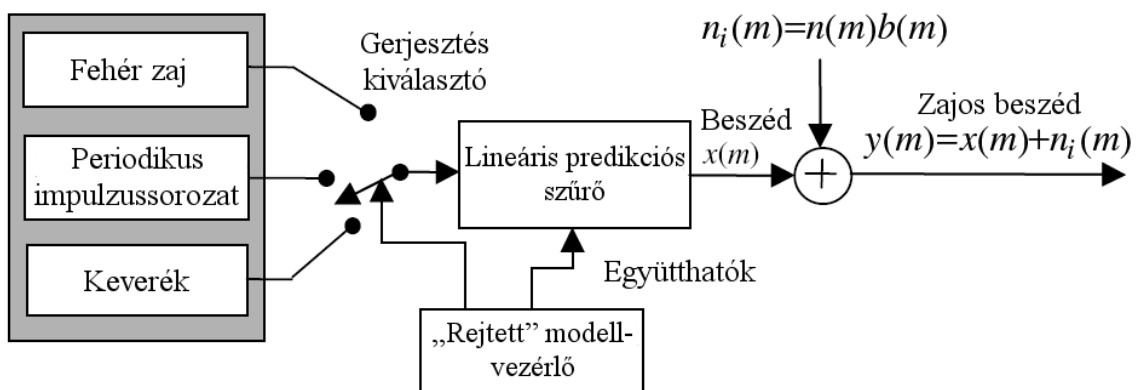
detektor megkeresi az impulzusok helyét, az interpolátor pedig kicseréli a torzult mintákat, figyelembe véve a torzult minta mindkét oldalán elhelyezkedő további mintákat. A detektor rész egy lineáris predikciós elemzőből, egy illeszkedésvizsgáló szűrőből és egy határérték keresőből áll. A kereső kimenete egy kétállású kapcsoló, mely az interpolátort vezérli. Impulzív zaj hiányában az interpolátor nincs bekapcsolva. Impulzív zaj jelenléte esetében az interpolátor bekapcsol és kicseréli a zaj által törölt mintákat.

## 3.3.4.3.1. Az impulzív zaj keresése

A legegyszerűbb módszer az ilyen jellegű zajok megtalálására, hogy meghatározzunk egy határértéket, és a határérték fölé eső mintákat zajnak osztályozzuk. Ez a módszer tökéletesen működik nagyon nagy impulzusok esetében, de ha a zaj amplitúdója a megadott határérték alá esik, akkor nem találja meg. A keresés eredményesebbé tehető, ha kihasználjuk a jel és az impulzív zaj jellemzőinek különbségét. Az impulzív zaj, vagy rövid időtartamú hullámvész bevisz egy a jelre nem jellemző megszakítást a jelbe. A megszakítás könnyebben megtalálható, ha a jelet szétválasztjuk. A szétválasztás (vagy digitális jelnél megkülönböztetés) művelete egyenértékű a kölcsönösségi viszony megszüntetésével, vagy a spektrális fehérítéssel. A lineáris prediktor modellezi a jel kölcsönösségi struktúráját, és ezen összefüggések megszüntetését végzi az inverz szűrő. A zajmentes  $x(m)$  jel lineáris predikciós modell segítségével a következő összefüggéssel írható fel:

$$x(m) = \sum_{k=1}^P a_k x(m-k) + e(m), \text{ ahol } a = [a_1, a_2, \dots, a_P]^T \text{ a lineáris prediktor } P \text{ rendű}$$

együttható vektora, és az  $e(m)$  gerjesztés egy zajhoz hasonló jel, vagy vegyes véletlenszerű zaj keveréke, és hullámvészok kvázi periodikus sorozata, melyet a következő ábra mutat. Az impulzív zajkereső nem a korrelálatlan kétállapotú impulzív zaj



30. ábra Zajos beszédmodell. A jelet egy lineáris prediktor modellezi, az impulzív zaj modellezésére pedig egy kétállapotú amplitúdómodulált modellt használunk.

alapján dolgozik, hanem azon a megfigyelésen alapszik, hogy a lineáris prediktorok jó modellt biztosítanak korrelált jelekre. A zajos  $y(m)$  jel a prediktor gerjesztett jelévé történő alakítása a következő hatásokat váltja ki:

- a.) Az jel amplitúdójának léptéke csökken az eredetileg gerjesztett jeléhez képest, mivel a zaj amplitúdójának léptéke változatlan marad, vagy növekszik.



- b.) A jel nem lesz kölcsönösen összefüggő, mert az impulzív zaj elmaszatólódott és átalakult léptékes alakba az inverz szűrő impulzus válaszaként.

Mindkét hatás javítja az impulzív zaj megtalálhatóságát. A beszéd és a zene hang időtartamának rezonanciája vagy a hangszer által erősített és spektrálisan alakított gerjesztésekből áll össze. Ez a gerjesztés véletlenszerűbb a beszédnél, és gyakran sokkal kisebb az amplitúdó-tartománya. Az impulzív zaj megtalálhatósága javítható az inverz szűréssel, mely lényeges lehet és függ a jel időben változó korrelációs struktúrájától. Ez a módszer egy optimális határértékképző használatával képes hatékonyan csökkenteni az impulzív zaj keresésének problémáját, a sorból kilógó minták megtalálását.

#### 3.3.4.3.2. Az impulzív zaj megtalálhatóság-javításának elemzése

A következőkben az impulzív zaj megtalálhatóságának javítását elemezzük ez inverz szűrő kimenetének vizsgálatával. Az előző egyenlet használatával a zajos jel modelljét a következő egyenlet segítségével (az előző átírt változata) írhatjuk fel:

$$y(m) = x(m) + n_i(m) = \sum_{k=1}^P a_k x(m-k) + e(m) + n_i(m), \text{ ahol } x(m), y(m) \text{ és } n_i(m) \text{ rendre a}$$

zajos jel, a jel és a zaj. Egy számítást használva  $\hat{\mathbf{a}}$  lesz  $\mathbf{a}$  prediktor együttható vektora, és a zajos  $y(m)$  jel inverz szűrhető és átalakítható  $v(m)$  zajosan gerjesztett

$$\text{jellé így: } v(m) = y(m) - \sum_{k=1}^P \hat{a}_k y(m-k) = x(m) - n_i(m) - \sum_{k=1}^P (a_k - \tilde{a}_k)[x(m-k) + n_i(m-k)],$$

ahol  $\tilde{a}_k$  a prediktor együttható számítási hibája. A zajos jel modelljének egyenlete ez alapján az egyenlet alapján következő egyenleté írható át egy másik alakba:

$$v(m) = e(m) + n_i(m) + \sum_{k=1}^P \tilde{a}_k x(m-k) - \sum_{k=1}^P \hat{a}_k n_i(m-k). \text{ Ebből az egyenletből három}$$

tétel határozható meg, mely elősegíti a zaj megtalálását a gerjesztési szekvenciában.

- a.) az  $n_i(m)$  impulzív zavar, mely általában domináns;
- b.) a régebbi  $P$  darab impulzív zajminta hatása az inverz szűrő működése végett, a jelenbe maszátolva:  $\sum \hat{a}_k n_i(m-k)$ ;
- c.) a gerjesztett jel varianciájának növekedése a paramétervektor számítási hibái miatt, mely a következő kifejezéssel írható le:  $\sum \tilde{a}_k x(m-k)$

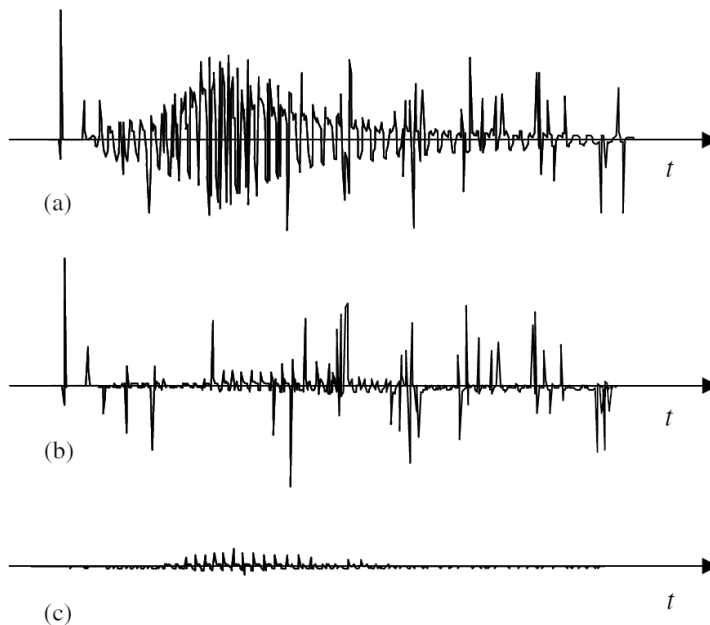
Az inverz szűrő eredményeinek javítása a következő formulákkal írható le. Az impulzív zaj a zajos jelhez viszonyított aránya az alábbiakban ismertetett módon

adható meg:  $\frac{\text{impulzív zaj energiája}}{\text{jel energiája}} = \frac{E[n_i^2(m)]}{E[x^2(m)]}$ , ahol  $E[\cdot]$  a várható érték operátora.

Az impulzív zaj detektálásában az érdeklődés tárgyát képező jel az impulzív zaj lesz, melyet a kísérő jelből kell megtalálni. Feltesszük, hogy a domináns zaj a  $v(m)$  zajos gerjesztési jelben az  $n_i(m)$  impulzív zaj, a két jel aránya a következőképpen írható fel:  $\frac{\text{impulzív zaj energiája}}{\text{gerjesztési energia}} = \frac{E[n_i^2(m)]}{E[e^2(m)]}$ . A teljes impulzív zaj jellel szembeni

aránya az előző két egyenlet hányadosából kapható meg, melyet a következő egyenlettel tudunk leírni:  $\frac{E[x^2(m)]}{E[e^2(m)]} = \text{nagyobbodás}$ .

Ez az egyszerű elemzés megmutatta, hogy az impulzív zaj megtalálhatósága függ a teljesítményerősítés karakterisztikájától, ennek következtében a lineáris predikciós modell rezonanciáitól. Beszéd esetében a jel amplitúdója zajmentes gerjesztett jelről lévén szó  $10^{-1}$  és  $10^{-4}$  között változik, ami a beszéd maga, ezért jelentős javulás érhető el az impulzív zaj megtalálhatóságában az inverz szűrő alkalmazásával a zajos beszédre. A következő ábra az impulzív zaj megtalálhatóságának



31. ábra Az inverz szűrő hatása a impulzív zaj megtalálhatóságára. (a) impulzív zajjal szennyezett beszéd, (b) az impulzív zajjal szennyezett beszéd gerjesztése, (c) beszéd gerjesztés impulzusmentesítve.

javítására alkalmazott inverz szűrő hatását mutatja. Az inverz szűrés hatása, hogy az  $x(m)$  jelet korrelálatlan  $e(m)$  jelgerjesztéssé alakítja, míg az impulzív zaj az inverz szűrő impulzív  $[1, -a_1, \dots, -a_p]$  zajátvitelének megfelelően torzul és megnövekszik, ahogy azt az iménti tétel mutatja a következő összefüggés alapján:

$\sum \hat{a}_k n_i(m-k)$ . Feltételezzük, hogy a gerjesztés egy

normális eloszlást követő fehér zaj, egy szűrő, melynek együtthatói az inverz szűrő együtthatóival egyeznek képes növelni a torzult impulzív zaj megtalálhatóságát a gerjesztési jelben.

## 3.3.4.3.3. Kétoldali prediktor az impulzív zaj megtalálására

Az impulzív zaj megtalálásának egyik hatékony módja, hogy a jelet korrelálatlanná tesszük. Ez a módszer felhasználható a kétoldali lineáris predikciós modellben is. Egy  $x(m)$  minta kétoldali lineáris predikciója  $P$  darab régebbi (múltbeli) és  $P$  darab jövőbeli minta alapján dolgozik, és a következő egyenlettel definiálható:

$$x(m) = \sum_{k=1}^P a_k x(m-k) + \sum_{k=1}^P a_{k+P} x(m+k) + e(m),$$

ahol  $a_k$  értékek a kétoldali prediktor együtthatói, és  $e(m)$  a gerjesztési jel. Minden elemzés arra az esetre készül, mikor csak egy oldalról vizsgáljuk a jelet, de ez kiterjeszthető a kétoldali modellre is. Azonban a gerjesztés mértéke a kétoldali modell esetében kisebb varianciát mutat, mint az egyoldali modellben, mivel az imént leírt egyenletben mind a múltbeli és mind a jövőbeli összes mintára modellezett a minták korrelációja. Habár az iménti egyenletben egy nem véletlenszerű szűrő van, az inverze szükséges a kereső alrendszerbe, mely már véletlenszerű. A kétoldali prediktor használata további javulást eredményezhet az impulzív zaj megtalálhatóságában.

## 3.3.4.3.4. Az eldobott minták interpolációja

Az impulzív zaj hatására a minták visszavonhatatlanul torzulnak, és eltávolításukkal „űr” keletkezik, melyet interpolációval próbálhatunk kitölteni. A hiányosságok interpolálására – ahhoz, hogy ne maradjon hallható nyomuk –, magas hanghűségű interpolátor szükséges. A legkisebb négyzetes autoregresszív (LSAR) módszer magas minőséget eredményez viszonylag kis számú az impulzív zaj miatt elveszett minta esetében. Az LSAR interpolációs módszer egy kétszintű folyamat. Az első szinten az impulzív zaj mindkét oldalán megmaradt minták alapján becsléssel meghatározzuk a jel predikciós modelljének paramétereit. A második szinten a meghatározott modellparaméterek és a hézag két oldalán található minták alapján interpoláljuk az elveszett mintákat. Ezzel az interpolációs módszerrel az impulzív zaj által eltorzított audiojelek magas minőségben visszaállíthatók.

## 4. Zajsűrés a gyakorlatban

### 4.1. Zenekari stúdiómunka

#### 4.1.1. A felvétel

Egy hangstúdióban legyen szó akár könnyűzenei vagy komolyzenei műről a hangrögzítés során számos problémával kell szembenéznünk. Amennyiben megoldható, az akusztikus (nem elektromos) hangszerek felvételét úgynevezett süket-szobában célszerű kivitelezni, a visszhanghatások elkerülése végett, melyek a hanganyag torzulását okozhatják. A felvételt adott esetben csatornákra bontva külön-külön rögzítjük. Ez azt jelenti, hogy az egyes hangszerek külön játsszák a darabot, majd a felvételt követően egy keverőpult segítségével születik meg a végleges hanganyag.

A felvétel során is felütik a fejüket a zajok. Amivel mindenképpen számolni kell, hogy a bemikrofonozott jelforrás valamilyen távolságra helyezkedik el a mikrofontól, néha a zenész átéléssel játszik, és a hangszerrel együtt mozog. A távolság és az alkalmazott mikrofon minőségének függvényében a rögzítő technika mindenképpen rögzít egy kevés környezeti zajt is. Ezen zajok ellen a környezet megfelelő megválasztásával lehet a legkönnyebben védekezni.

További problémákat okoznak a mikrofonok viszonylag alacsony jelszintjei, és a jeleket továbbító kábelek. A kábeleket célszerű lehetőség szerint minél vastagabbra, tiszta anyagból készültre, és minél rövidebbre választani, a lehető legkevesebb csatlakozóponttal, ugyanis ez tudja a jel legjobb minőségű szállítását elvégezni. A nagy keresztmetszet és a viszonylag kis hossz biztosítja a kis ellenállást, mely a jelveszteséget csökkenti, és a tiszta anyag (például oxigénmentes réz, négy kilencses tisztasággal – 99,99 % réz) pedig a szennyezések miatt kialakuló örvényáramok okozta zajtól óvnak bennünket. A csatlakozópontok további torzítást és zajt eredményezhetnek, az anyagok különbözősége, és a csatlakozó felületek mérete miatt. További eredményjavulás várható az arany, vagy felületén aranyozott kábelek alkalmazásától a magas frekvenciájú jelek átvitele szempontjából – a skin effektus miatt. A skin effektus azt jelenti, hogy a váltakozó áramot szállító vezetékben az elemi töltéshordozók a vezetőben annál távolabb haladnak a vezető középpontjától, minél nagyobb az átvitt jel frekvenciája.

Az alacsony jelszint miatt a digitális feldolgozhatóság elérése érdekében előerősítő fokozatokat kell a rögzítő készülék és a jelforrás(ok) közé építeni, melyek esetében az alkatrészek kialakítástól és minőségétől függően termikus zajgenerátorként működnek. A sávátviteli problémákat többnyire nem az erősítőelemek okozzák, hanem a mikrofonok.

További érdekesség a jelforrás szempontjából, hogy az elektromos hangszereket, mint például egy elektromos gitár, vagy basszusgitár általában egy úgynevezett alap erősítőre csatlakoztatjuk, majd ezen erősítő hangszórója elé helyezünk el mikrofont. Általában e hangszerek esetében az effekteket már az erősítő előtt, vagy annak effektútjába csatlakoztatjuk. Ilyenkor egy torzító effekt képes a hangszer hangszedője (pickupja) által felfogott zajokat nagymértékben felerősíteni, amit még – egy a torzító előtt elhelyezett (a játéktechnikát könnyítő) – dinamika kompresszorral lehet „növelni”. Ezek a zajok elektromágneses zajok, és sajnos a hangszer működéséből adódóan nehéz ellenük védekezni, mivel a húrok rezgését is ilyen módon érzékeli a hangszedő. Ilyen hangszerek esetén már két termikus zajgenerátor rendszer kapcsolódik a rögzítő elé, ezzel fokozva a zajszennyezést.

A rögzítő technika manapság már többnyire digitális rendszer, melynek az érkező jeleket egy ADC segítségével digitálissá kell alakítania. A jel ekkorra már zajjal erősen szennyezett lesz, a korábbi berendezések felépítésének és minőségének függvényében. A digitalizálás a mai korszerű hangkártyáknak nem jelentenek problémát, mégis stúdió célokra külön gyártanak – magasabb árfekvésű, jobb minőségű – kártyákat. A mintavételezési frekvencia megválasztása ekkor az egyik fő probléma. Általában a végeredmény CD lemezre kerül, ahol szabvány írja elő a 44100 Hz-es mintavételezési frekvenciát. Vannak hangkártyák, melyek ilyen mintavételezési frekvencia mellett nem produkálnak lineáris sávátvitelt, ellenben 48000 Hz-en igen. A felvett jelet visszahallgatva, van különbség 48000 Hz-en és a 44100 Hz-en mintavételezett jelek között az első javára, bár azt gondolhatnánk, hogy a hallható tartományon már ezek a frekvenciák kívül esnek. Másik érdekes kérdés a kvantálási hossz. A mai világban már a megfelelő dinamika elérése érdekében 16 bitnél nagyobb értéket választunk.

#### 4.1.2. Az utómunka

Az utómunkálatok első lépése a felvétel során a jelre került zajok eltávolítása. Ez esetben, ha sávonként rögzítettünk, akkor minden sávra külön célszerű a zajt eltávolítani, egészen pontosan annak hatását csökkenteni, ugyanis a zajok vélet-

lenszerű mivolta nem teszi lehetővé teljes eltávolításukat. Az ilyen felvételek esetén azért kell külön kezelni a zajokat, mert egy mikrofon és egy elektromágneses hangszedő több zajgenerátorral együttműködve más és más jellegű zajokat eredményeznek. Ha például egy dobfelszerelésre gondolunk, akkor észrevehető, hogy a megfelelő minőségű jel érdekében célszerű a hangszer egyes darabjait külön be-mikrofonozni, adott esetben más típusú mikrofonokkal, tehát ekkor már egy mikrofonrendszerrel beszélünk. Az egyes mikrofonok hangját külön kell bontani hangszín szempontjából, esetleg sztereó felvétel esetében még a panoráma sem fog egyezni minden esetben (például: lábcintányér jobboldal, pergődob közép, kísérő cintányér baloldal, természetesen a zenész kezességét figyelembe véve lehet fordítva is). A hangszínt minden esetben az egyes részeken külön kell beállítani, ami függ a mikrofon minőségétől, távolságától, helyzetétől és állásától. Ez a mikrofonrendszer mikrofononkénti hangszín beállításokkal másfajta zajt állít elő, mint például az imént említett elektromos gitár hangszedője egy torzítóval. Ezért is fontos a sávonkénti felvétel. Ezek a zajok csak ritka esetben tartalmaznak impulzív jellegű zajokat (vagy ha az ilyen zajok megjelennek, azok megszüntetéséről még a felvétel előtt célszerű gondoskodni), ezért a zaj csökkentésére a legalkalmasabb módszer a spektrális kivonás. Fontos, hogy a zajszűrő program (algoritmus) olyan legyen, hogy a hangképet ne, vagy csak nagyon kis mértékben módosítsa, a legjobb eredmény elérése érdekében.

Ezt követően az egyes hangszerek még további effektek hozzáadásával színesíthetők, valamint itt nyílik lehetőség az utolsó hangszínbeállításra, ugyanis a felvételkor még csak nyersanyag keletkezik. Itt adott esetben az egyes sávokat további analóg keverőpultba küldhetjük, de léteznek olyan stúdióban alkalmazható programok, melyek képesek a többsávós felvétel kezelésére. Ilyen program például a korábbi (a 2003-ban az Adobe cég által megvásárolt Syntrillium cég fejlesztésében megjelent) Cool Edit, és a manapság helyette alkalmazott Adobe Audition program, melyek kereskedelmi forgalomban kaphatók, de van ingyenes időkorlátos próbaváltozatuk is. Ezek a programok nem csak többsávós hanganyag effektezésére, hangszínbeállítására alkalmasak, hanem az azt megelőző zajcsökkentésekre is.

Természetesen a megfelelő hangzás elérése érdekében a hanganyagot a hangfelvétel elkészülte után hetekkel célszerű keverni, majd a végeredményt többféle berendezésen visszahallgatni, a maximális minőség elérése érdekében. A megfelelő hangzás elérése után még egy dinamika kompresszor segítségével lehet az anyagot

„összetömöríteni”. A mintavételezési frekvencia és a kvantálási hossz a későbbi hanghordozó függvényében változhat a rögzítési értékekhez képest, de ez sajnos a jel torzulását és néha még további zajok megjelenését is okozhatja.

## **4.2. Hagyományos hanglemez, hangkazetta archiválása CD-re**

Minden archiválási folyamat az adott hanghordozón tártolt hanganyag digitalizálásával kezdődik. Ezt akár otthonunkban is megtehetjük, feltéve, hogy rendelkezünk a forrás hanghordozó megfelelő minőségű lejátszására alkalmas berendezéssel. A hangminőséget jelentősen befolyásolhatja a jelforrás, továbbá az is, hogy milyen eszközzel végezzük a digitalizálást, és hogy milyen kábelek kötik össze az adott berendezéseket. Itt is lehet környezeti elektromágneses zavarokra számítani.

### *4.2.1. Hanglemez archiválása*

#### 4.2.1.1. A lejátszó rendszer

A gyártók általában többféle megoldást alkalmaznak a lemezejátszó mechanikájának és a mechanikai mozgásokat elektromos jellé alakító részegységek kivitelezésére.

A mechanika leglényegesebb elemei:

- Lemeztányér, amely a lemez forgatásáért, megtartásáért felelős, továbbá sok esetben a hangkar mozgatásával is foglalkozik, az alsó mechanikus szerkezeten keresztül. Többnyire nagy tömegű, az egyenletes forgás megvalósítása érdekében (nagyobb tehetetlenség, így nem képes a motor gyors fordulatszám-változásainak követésére). Fontos a jó csapágyazás és kenés a könnyű futás érdekében.
- Motor, amely a lemeztányért hajtja, nagyon fontos, hogy állandó fordulatszámú legyen nagyon kicsi ingadozással, mely valamilyen általában erőzárás közlőművel kapcsolódik a lemeztányérhoz (ez esetben fordulatszáma sokkal nagyobb, mint a tányéré).
- Hangkar (esetleg a mozgató mechanikával együtt), mely hangszedő fejjel együtt a lemeztűt egy shellnek nevezett tartóban rögzíti, és gondoskodik a megfelelő tűnyomásról és a barázdakövetésről.

A lemeztányér meghajtására háromféle megoldást szoktak alkalmazni, az egyik a dörzshajtás, a másik a szíjhajtás – gyakorlatilag ez az egyik legelterjedtebb –, és végül a direkt hajtás – ez közvetlen a motor forgórészére rögzített viszonylag

könnyű lemeztányért jelent. A motor lehet váltakozó áramú, mely többnyire a szabványos hálózatban alkalmazott frekvenciára épít a forgás tekintetében, vagyis szinkron motor, és lehet úgynevezett DC (egyenáramú) motor is. A közvetlen hajtású lemezjátszók esetében a legpontosabb fordulatszám elérése végett kvarcvezérlést alkalmaznak. Egyes típusoknál egy stroboszkóp lámpa segíti a fordulatszám beállítását, egy potenciométer beiktatásával, melyet általában az egyenáramú motorral szerelt típusok esetében alkalmaznak. A három megoldás közül a kvarcvezérlésű direkt hajtás a legpontosabb megoldás.

A hangkart mechanikusan áthelyezni képes lemezjátszók a lemeztányér alatt elhelyezett mechanikus szerkezet segítségével teszik mindezt. Létezik belőlük fél-automata és automata kivitel is. A hangkarban helyet foglaló (sok esetben leszerelhető, cserélhető) shell biztosítja a lemeztű számára kialakított hangszedő fej rögzítését. A szögkiegyenlítés végett – nem gyakran ugyan – alkalmaznak úgynevezett tangenciális hangkart, melynél nincs ilyen jellegű probléma. A nem tangenciális esetben a fej valamekkora szöget zár be a hangkarral, melyet vagy a hangkar meghajlításával, vagy a shell kialakításával érnek el.

Hangszedőből háromféle alaptípus létezik, melyek működésükben különböznek egymástól.

Az egyik – régebben viszonylag olcsó és hatékony megoldást biztosító – típus a kristály hangszedő. Ez működését tekintve a gyémánt vagy zafír tű által a lemeztől letapogatott rezgéseket egy vékony cső segítségével egy piezokristályra vezeti, mely a rezgést elektromos jellé alakítja. Ez a letapogató rendszer viszonylag nagy túnyomással működik jól, és ez sajnos nem kíméli a mikrobarázdás lemezünket. Hangzás tekintetében nem túl jó, de kielégítő eredmény elérhető vele. Nagy előnye, hogy nem igényel külön korrekciós előerősítő fokozatot, viszont impedancia-illesztés szempontjából nagyimpedanciás bemenetet igen.

A másik szinte a legelterjedtebb, már jó hangzást biztosító megoldás a mozgó mágneses hangszedő. Ez esetben (sztereó kialakítást feltételezve) két egymással 90°-os szöget bezáró tekerccs között mozgatja a vékony cső végére rögzített mágnest a lemeztől letapogatott mechanikai rezgés. Túnyomás tekintetében változatos megoldásokkal találkozhatunk, mert van itt  $\frac{3}{4}$  grammtól kezdődően kristály hangszedőt megszegyenítő 5 grammos nyomást igénylő kialakítás is. Itt már kell egy – a köznapi életben alkalmazott erősítők nagytöbbségében az alkalmazás idő intervaluma alatt fellelhető – korrekciós előerősítő fokozat, mely a mély hangokat kiemeli, a közép hangokkal nem tesz semmit, és a magas hangokat pedig csillapítja. Erre



azért van szükség, mert a lemez előállításakor így rögzítik a jelet – a barázda szélességének csökkentése miatt (több hanganyag fér el egy lemezen). Viszonylag egyszerűen cserélhető az elkopott tű, melyet még manapság is forgalmaznak bizonyos helyeken.

A harmadik megoldás a mozgó tekercses megoldás. Ez csak fejjel együtt cserélhető, nagyon drága, és hangzásvilágát tekintve ez a legjobb minőségű. Az előerősítő fokozat ide is kell, de az erősítésének nagyobbak kell lennie az előzónél, mert a jel feszültsége csupán kb. az egytizede a mozgó mágneses hangszedőének. Létezik olyan erősítő, amely mindkét jelet képes fogadni. A megoldások közül a legjobb hangminőség a mozgó tekercses hangszedővel érhető el, de ára, és a szükséges további eszközök miatt inkább a mozgó mágneses hangszedőt használjuk.

#### 4.2.1.2. Várható zajok

A hanglemezek mechanikus adathordozók, ezért nagyon sérülékenyek. A nem megfelelő tárolás következtében a por komoly károkat tud okozni a hangzásban. A nem megfelelő poreltávolítás apró karcokat okoz, melyek hallhatók lesznek lejátszáskor. A lemez felületére az ujjainkról került zsír is károsítja a hanghordozót, továbbá a felületen nagyobb karcok is kialakulhatnak (a nem megfelelően mozgatott hangkar következtében a tű is okozhat ilyen karcokat), melyek az impulzív zajokért felelősek.

A hanglemez kialakítása során a felületi egyenetlenségek keletkeznek, melyek szintén hallhatóvá válhatnak lejátszáskor. Ezért van különbség a keleti és a nyugati nyomású lemezek között, mivel más a gyártás alapanyaga és némileg a technológiája is.

A lemezjátszó mechanikája is képes zajokat előállítani. A lemeztányér a csapágyazás és kenés függvényében valamilyen mértékben rezeg. Ez egy precízen csapágyazott motor esetén egy direkt hajtású lemezjátszóban nem okoz problémát, mert általában a hallható tartományon kívül esik a rezgés frekvenciája.

Különbséget kell tenni a lemez száraz, illetve nedves lejátszása között is. A nedves lejátszás esetén a lemezt mosófolyadékkal lejátszás közben mossuk, így a szennyeződések eltávolítása is megtörténik, továbbá a folyadék vékony réteget képez a tű és a barázda információt tartalmazó részei között, ezért a hangzás is változhat nagyon kis mértékben.

A lemez kivitelezésénél fogva a korrekciós előerősítő fokozat is zajt termel, majd az azt követő illesztő fokozat viselkedik zajgenerátorként. Ezekben az erősí-

tőkben az aktív elemek melegszenek, termikus zajt előállítva. Továbbá a nagy erősítés és a mágneses működés miatt érzékenyek az elektromágneses zajokra is, mely megfelelő szigeteléssel orvosolható.

Tehát a lemezen megjelenik a helyreállítandó hézagokat maga után vonó impulzív zaj, és a spektrumát tekintve nemfehér zaj is, melyeket két különböző eljárással tudunk csökkenteni.

#### 4.2.1.3. Zajcsökkentési eljárás

A felvétel elkészítésekor figyelni kell az alkalmazott technikák jelszint és impedancia-illesztésére, valamint a mintavételezési frekvenciára és a kvantálási hossza is. Lehetőleg a cél adathordozóéval kompatibilis értékeket kell választani, a minőségromlás elkerülése érdekében.

A lemezjátszók készülhettek a keleti, illetve nyugati régióban is. A keleti régió általában az ötpólusú tuchel aljzatokat alkalmazta a jelek továbbítására, amely nem csak kialakításában, hanem jelszintjében és impedanciájában is eltér a nyugati régióban elterjedt (4 darab) RCA csatlakozókhöz képest. A tuchel csatlakozók kisebb jelszinten dolgoznak, és nagyobb bemeneti impedanciát igényelnek, továbbá kisebb felületen érintkeznek, tehát nem alkalmasak közvetlenül digitalizálásra. (Természetesen egy egyszerű erősítőfokozat közbeiktatásával alkalmassá tehetők a feladat elvégzésére az ilyen berendezések is.) A célszerű választás tehát a folyamat kivitelezésére egy *nyugati gyártású lemezjátszó, mozgó mágneses* hangszedővel, a megfelelő előerősítő fokozatot tartalmazó szintén *nyugati gyártású erősítővel*, mely illeszthető a szabványos hangkártyák vonalbemenetére.

A folyamat a digitalizálással kezdődik. A feladatra a számítógépbe célszerű egy eleve jó értékű jel–zaj aránnyal rendelkező, kis torzítású hangkártyát választani. A Creative cég régebben is készített ilyeneket, és ma is gyárt alkalmas eszközöket. (A Sound Blaster sorozat elemei.) A digitalizáláskor a jel erős zajszenyezéssel érkezik az eszközbe.

Az első lépés az impulzív zajok megkeresése és eltávolítása. Erre alkalmas az előző részben említett Adobe Audition program. Itt a határértékek feltérképezése után indulhat a folyamat, sajnos az eredmény csak a végrehajtás után lesz hallgatható, nincs „előzetes bemutatás” lehetőség. A folyamat az impulzív zaj mennyiségétől és a számítógép sebességétől függően akár több órát is igénybe vehet. Közben látható, hogy hány mintát vélt zajnak, és hány mintát sikerült kijavítania a programnak. Természetesen az javított/kihagyott arány a beállítás függvénye, na-

gyon karcos lemez esetén sok impulzív zaj van, melyek javítása eredményezheti a hangminőség romlását is, oly formában, hogy a magas frekvenciájú jeleket is impulzív zajnak tekintik.

A következő lépés a spektrális kivonás. Ekkor mintát kell venni egy csak zajos részből, mely lehet a lemez bevezető része. A mintában szereplő pillanatképek száma beállítható, de nem feltétlenül eredményez jobb minőséget a nagyobb szám. Fontos paraméter a zaj csökkentésének mértéke, a precizitási szorzó és a simítás is. Ezek együttesen határozzák meg az eredmény minőségét. Van „előzetes bemutatás”, tehát meghallgathatjuk, hogy milyen eredményre számíthatunk, sőt megfigyelhető csak a zaj is.

A végén célszerű a jelet normalizálni, a hangerő (vagy jel–zaj arány) javítása érdekében. A mentés történhet szakaszonként is, ezzel a hanganyag darabolása is megtörténhet. Az eljárással akár 60–80 dB jel–zaj arány is elérhető, ami a kiindulási 30–40 dB-hez képest nem rossz eredmény.

#### 4.2.2. Hangkazetta archiválása

##### 4.2.2.1. A magnetofon és a hangszalag

A magnetofon vékony mágnesezhető réteggel bevont műanyag szalagon rögzíti a jeleket. A jelek, mivel az átmagnesezhető réteg vékony, nagyon kis feszültséget indukálnak a letapogató rendszerben, vagyis a magnetofon fejben, mely egy speciális kemény, kopásálló bevonattal rendelkező tekercs(pár). A kicsiny indukált feszültség erősítése egy korrekciós előerősítő fokozattal kezdődik, melynek a magas hangokat kell kiemelnie. A lejátszó rendszer ezt a szalagot húzza egyenletes sebességgel a fej előtt. A szalagsebességről egy kisebb átmérőjű fém és egy rugóval rászorított nagyobb átmérőjű gumigörgő gondoskodik, egyes esetekben ezt a rendszert megduplázzák, a stabilabb szalagsebesség érdekében. A görgők viszonylag könnyen szennyeződnek, ami a szalag nem kívánt oldalirányú mozgásához vezet; ez károsan befolyásolja a lejátszást. A szalagsebesség egy fontos jellemzője a berendezésnek, ami 4,76 cm/s egy átlagos magnetofon esetében. Ennek beállítása referenciaszalaggal és oszcilloszkóppal történik. A gyártók egy túrésen belülre állítják ezt az értéket, tehát két azonos szalagsebességű berendezést nehéz találni. A lejátszó és felvevő kombinált fej szennyeződhet, átmagneseződhet, ami szintén minőségromlást eredményezhet. Fontos paraméter még a szalagon található sáv és a fejben található érzékelő egymáshoz viszonyított elhelyezése is. Nem megfelelő beállítással a magasabb frekvenciájú hangok mennyiségének nagymértékű csökkenése

érhető el. A gyártók ezt is határértékek közé állítják, de létezik gyárilag rosszul beállított magnetofon is.

A szalag esetén problémák a szalagra felvitt mágnesezhető réteg minőségével kezdődnek, ebből alapjában véve négyféle típust alkalmaztak. Az első a normál vasoxid, a második a króm dioxid, a harmadik a ferrokróm (ezt már nem alkalmazzák) a negyedik pedig a magnetit. A következő probléma a mágnesezhető réteg szemcsézete, egészen pontosan a mágneses információhordozók elrendezése. Ez a szalag gyártójától függően – azonos típusú mágnesezhető anyag esetén – anyagminőségben is eltérhet egymástól, tehát vannak jobb minőségű és rosszabb minőségű hangkazetták, annak függvényében, hogy a mágnesességet mennyi ideig képesek megőrizni az egyes szalagok. A harmadik probléma a műanyag réteg vastagsága, mely lejátszáskor is okozhat problémát a szalagsebesség és az oldalirányú mozgás miatt, de tárolás esetén kérdéses, hogy az egyes átmágnesezett rétegek hogyan hatnak egymásra az idő múlásával. A műanyag hosszirányban képes a megnyúlásra, ami tovább fokozza a problémákat. Továbbá a vékony szalag nagyon érzékeny a mechanikai behatásokra, mint például a gyűrődés, amely durva minőségromláshoz vezethet. Az ilyen jellegű hiányok sem pótolhatók a jel folytonosságában, még digitális módszerrel sem. Nagyon fontos szerepe van még szalagok esetén az előmágnesezésnek, ugyanis ez is potenciális zajforrás.

#### 4.2.2.2. Várható zajok

A szalagos egységek esetében az előmágnesezési zaj után a szalagnak is van egy meghatározott értékű zaja, melyet egy olyan elektronika követ, ahol több termikus zajgenerátor is működik egy időben. A zaj ezért összetett, és általában egy elég jól hallható sistergés formájában jelentkezik, tehát főleg a magasabb frekvenciákon ölt jelentős mértéket.

A magnetofon esetén alkalmazható a Dolby zajcsökkentő rendszer, de lejátszáskor csak akkor, ha a felvétel is azzal készült, ellenkező esetben nem csak a zaj, hanem a magas hangok is eltűnnek. Ha nem tudjuk egy szalag esetén, hogy használtak-e valamilyen analóg zajcsökkentőt a felvételkor, akkor a berendezés zajcsökkentőjét kapcsoljuk ki, mert így a digitális zajszűrő hatékonyabban távolíthatja el a zajt, és több hasznos jel maradhat épen.

A magnetofonnal – a gyártók ígéretei szerint – akár 60 dB jel–zaj arány is elérhető a Dolby zajszűrővel, de sajnos a szalagok minősége és a közbeiktatott erősítőfokozatok miatt a digitalizálás után ez az érték 40–45 dB-re romlik.

Mivel vannak olyan rögzítő eszközök ebben a kategóriában, amelyek a törlést és egyben az előmágnesezést egy állandó mágnessel valósítják meg, ezért az előmágnesezés szintje közel sem lesz állandó, sokkal inkább jól hallható alacsony frekvenciájú előmágnesezési zaj lesz belőle. Ez a zajforma oly mértékben elnyújtott, és véletlenszerű, hogy algoritmikus csökkentése esetenként nem is lehetséges, ugyanis a zaj mintavételezése során nem választható ki úgy zajos rész, hogy e zajból kellő mennyiségű és minőségű jusson a mintába megfelelő kivonáshoz.

#### 4.2.2.3. Zajcsökkentési eljárás

A magnetofonszalagokról érkező jelek főként magas frekvenciájú zajt tartalmaznak, mely a korrekciós előerősítő fokozatból adódik. Ez a zaj mintavételezést követően spektrális kivonással csökkenthető a zavaró mérték alá.

Csak olyan esetekben ajánlott a magnetofonszalagok digitális hanghordozóra másolása, ha az anyag bír akkora fontossággal, hogy megtegyük (más hanghordozón, nem hozzáférhető). A magnetofonszalagokról tudjuk, hogy a normál szalag sávátvitele alig több mint a fele a krómdioxidos vagy magnetit szalagokéhoz képest, és a szalag zaja, ami a mágneses információhordozó minőségéből adódik, is lényegesen nagyobb az ilyen szalagoknál.

A magnetofonnál általában nem jelentkezik impulzív jellegű zaj, ezért a zajcsökkentés elvégezhető egyetlen lépésben. A digitalizált jelből (az elejéről vagy a végéről) zajmintát veszünk, majd pedig spektrális kivonással távolítjuk el a zajt. Az eredmény – bár az eljárás egy lépésből áll csupán – nem lesz olyan minőségű, mint a hanglemezek esetében, ugyanis a magasabb frekvenciájú hangokat a rögzítő rendszer és a szalag már korábban eltüntette, ezért lejátszáskor sem voltak már jelen. Ezek a hangok már semmilyen eszközzel nem állíthatók vissza.

## 5. Összegzés

Dolgozatom célja volt a környezetünkben előforduló zajok osztályozása, és azok csökkentésére megoldást találni. A potenciális zajforrások nagytöbbségére a dolgozatom első szakaszában világítottam rá. Tekintettel arra, hogy zaj hasonlóan az információt hordozó jelhez véletlenszerű, ezért modellezése és csökkentése közel sem olyan egyszerű, mint ahogyan azt a köznapi ember gondolja. A zajok modellezése a zajcsökkentés szempontjából jelentős, ugyanis egy megfelelő modell segítségével az eljárás könnyebben algoritmizálható, és alkalmazható a digitális jelfeldolgozásban. A számítógépek miniatürizálásával és gyors fejlődésével korunkban egyre nagyobb teret kap a kommunikációs és jelrögzítő rendszerekben a digitális jelek alkalmazása, mely a feldolgozhatóság és a hozzáadódó zajokkal szembeni jó „ellenálló-képességének” köszönheti sikerét.

A zajok modellezését kétféle szempont szerint közelítettem meg dolgozatomban, melyekből kiderült, hogy a véletlenszerű jelek előre meg nem jósolható értékeket is felvehetnek. A modellek alkalmazhatók más köznapi véletlenszerű események modellezésére is, melyek messze esnek az informatika tudományterületétől, mint például a közgazdaságtan, vagy a meteorológia, vagy az orvostudomány.

A probléma megoldása a zajcsökkentési eljárások alkalmazása. Többféle módszer létezik, melyekből csak néhányat volt alkalmam kiemelni, dolgozatom szűkös keretei miatt. A módszereket az adott problémának megfelelően, jelen esetben a zajok fajtáihoz illően kell alkalmazni, a várt eredmény elérése érdekében. A zajok csoportosítása, és forrásuk meghatározása ezért nagyon fontos, a legalkalmasabb eljárás kiválasztása szempontjából. Egy véletlenszerű jelsorozatból bizonyos esetben nehéz eldönteni még a szétválogatás után is, hogy melyik a számunkra hasznos információt hordozó jel. Erre nagyon szemléletes példa egy gépjármű motorja, mely az utas szempontjából zajforrás, de egy hozzáértő szerelő sok értékes információhoz juthat a hallott hangokból.

Az utolsó szakaszban az elméleti ismeretek gyakorlatba ültetését próbáltam kifejteni, saját tapasztalataim alapján. Több mint tíz évvel ezelőtt volt alkalmam több „házi stúdióban” készült könnyűzenei demófelvételhez segítséget nyújtani, és jómagam is aktív részese voltam egyes felvételek elkészítésének, mint hangszeres „zenész”. Jelenlegi iskolámban néha alkalmam nyílik kamatoztatni korábbi ismer-

reteimet, és dolgozatom megírásával nagyobb rálátást nyerhettem a későbbi munkámhoz a zajok forrásaira és megfelelő mértékű csökkentésükre.

A régebbi hanghordozókon tárolt, többnyire zenei felvételek korszerű adathordozóra mentése több ismerősömnek eszébe jutott már, de megfelelő eszköz hiányában maguk nem tudták elvégezni a műveletet. Érdeklődési körömbé tartozik a hangrendszerek felépítésének, és működésének megismerése és tanulmányozása, ezért több, mikrobarázdás lemez lejátszására alkalmas eszköz és magnetofon is található otthonunkban, mellyel a hanganyagok digitalizálásához forrást tudok biztosítani. A számítógépben található hangkártyákra mindig figyeltem, hogy ne csak gyenge minőségű hangjelet tudjon produkálni a komolyabb hangrendszer számára, hanem megfelelő működésével minőségi zenehallgatásra legyen alkalmas.

A dolgozat megírása során törekedtem arra, hogy a címben megragadott témát minél jobban, és több szempont szerint is kifejthessem, amit végeredményében sikerült megvalósítanom.

## 6. Irodalomjegyzék

1. Saaed V. Vaseghi – Advanced Digital Signal Processing and Noise Reduction  
Second Edition  
Copyright © 2000 John Wiley & Sons Ltd.
2. Valószínűségszámítás és statisztika előadások anyaga  
Nagy Márta – 2001–2002 tanév
3. Molnár-Sáska Gábor – Rejtett Markov modellek statisztikai vizsgálata  
Tézisfüzet – 2005  
forrás: [http://doktori.math.bme.hu/Molnar-Saska\\_tezisfuzet.pdf](http://doktori.math.bme.hu/Molnar-Saska_tezisfuzet.pdf)
4. dr. Tolnai János – Rádióamatőr vizsgára felkészítő tananyag  
Puskás Tivadar Távközlési Technikum, 2006.  
forrás: [http://www.puskas.hu/r\\_tanfolyam/r\\_tananyag.html](http://www.puskas.hu/r_tanfolyam/r_tananyag.html)



Debreceni Egyetem  
Informatika Kar

# Szakedolgozat

## Zajok és eltávolításuk

**Témavezető:** Dr. Papp Zoltán  
egyetemi adjunktus

**Készítette:** Gambár Péter István  
III. informatika tanár szakos,  
levelező tagozatos hallgató

Debrecen  
2009.

# Tartalomjegyzék

1. Bevezetés.....	4
1.1. A cím magyarázata.....	4
1.2. Témaválasztás indoklása .....	4
2. Zajok és modellezésük .....	5
2.1. Zajok fajtái és forrásai.....	5
2.1.1. Általános felosztás .....	5
2.1.2. Fehér zaj.....	7
2.1.3. Színezett zaj .....	7
2.1.4. Impulzív zaj.....	8
2.1.5. Tranziens zaj impulzus.....	10
2.1.6. Termikus zaj .....	11
2.1.7. Schottky-zaj.....	12
2.1.8. Elektromágneses zaj .....	13
2.1.9. Csatornatorzítások.....	14
2.1.10. Zajmodellezés .....	15
2.1.10.1. Hozzáadódó fehér normális eloszlású zajmodell (AWGN).....	16
2.1.10.2. Rejtett Markov modell.....	16
2.2. Zajmodellek .....	18
2.2.1. Valószínűségi modellek.....	18
2.2.1.1. Véletlenszerű jelek és sztochasztikus folyamatok.....	18
2.2.1.2. Valószínűségi modellek .....	19
2.2.1.3. Állandó és nem állandó véletlen folyamatok .....	23
2.2.1.4. Véletlen folyamatok várható értéke.....	24
2.2.2. A rejtett Markov modell.....	29
2.2.2.1. Nem stacionárius folyamatok statisztikai modellje .....	29
2.2.2.2. Rejtett Markov modellek.....	31
2.2.2.3. A rejtett Markov modell betanítása .....	34
3. Zajsűrítési módszerek.....	36
3.1. Zajzár .....	36
3.2. Analóg valós idejű zajsűrítők.....	36
3.2.1. Dolby zajsűrítő.....	36
3.2.2. Zajérvénytelenítés.....	38
3.3. Digitális jelek feldolgozása.....	39

---

3.3.1. Átalakítások .....	39
3.3.1.1. Analóg–digitális átalakítás .....	39
3.3.1.2. Digitális–analóg átalakítás .....	40
3.3.2. A Wiener szűrő.....	42
3.3.2.1. A legkisebb négyzetek becslése .....	42
3.3.3. Zajcsökkentés spektrális kivonással .....	44
3.3.3.1. A spektrális kivonás módszere.....	45
3.3.3.2. A spektrális kivonó megvalósítása.....	47
3.3.4. Az impulzív zaj eltávolítása.....	50
3.3.4.1. Impulzív zaj matematikai modellje .....	50
3.3.4.2. A medián szűrő.....	51
3.3.4.3. Impulzív zaj eltávolítása lineáris prediktív modell segítségével ...	53
4. Zajszűrés a gyakorlatban.....	58
4.1. Zenekari stúdiómunka .....	58
4.1.1. A felvétel .....	58
4.1.2. Az utómunka.....	59
4.2. Hagyományos hanglemez, hangkazetta archiválása CD-re.....	61
4.2.1. Hanglemez archiválása.....	61
4.2.1.1. A lejátszó rendszer .....	61
4.2.1.2. Várható zajok .....	63
4.2.1.3. Zajcsökkentési eljárás .....	64
4.2.2. Hangkazetta archiválása.....	65
4.2.2.1. A magnetofon és a hangszalag.....	65
4.2.2.2. Várható zajok .....	66
4.2.2.3. Zajcsökkentési eljárás .....	67
5. Összegzés .....	68
6. Irodalomjegyzék.....	70

## 1. Bevezetés

### *1.1. A cím magyarázata*

Általában hangfelvételek (függetlenül attól, hogy milyen eszközzel készülnek és milyen adathordozóra) és kommunikációs rendszerek esetében felütik fejüket az additív zajok. Ezen zajokat a hasznos jelhez szokták viszonyítani, és sajnos adott esetben a hasznos jel nagyságát is meghaladhatják, mely nagyban károsítja a jelek értelmezhetőségét. Több esetben azonban lehetőség nyílik az ilyen zajok mértékének, erősségének csökkentésére, mind hagyományos, mind digitális rendszerben. A zajok eltávolításánál nagyon fontos, hogy a hasznos jel minél nagyobb hányada maradjon ép, értelmezhető, és lehetőleg torzításmentes. Zajforrástól függően a zajok csökkentése hagyományos úgynevezett analóg rendszerekben is megvalósítható, akár megfelelő alkatrészek és eszközök alkalmazásával is. Jelen dolgozatban a tárolt és szállított hangjelek zajainak felismerésével, fajtaival és azok csökkentésével foglalkozom.

### *1.2. Témaválasztás indoklása*

Több éve egy zenei általános iskolában tanítok, ahol esetenként szükség van a kórus munkájának megörökítésére nem csak képi, de hang formájában is. A hangfelvételek megfelelő minőségű elkészítése és tárolása érdekében a rögzítő technika és az előtte található átalakító rendszerek által termelt zajok eltüntetése adott esetben nagyon fontos. Céлом, hogy a dolgozat megírásával iskolámban minél hatékonyabb munkát tudjak végezni ebben a tekintetben is.

Személyes érdeklődési körömbé tartozik a zenei hangfelvételek digitális feldolgozása, illetve az erre alkalmas berendezések és szoftverek megismerése is, tehát többek között ez is motivált a szakdolgozat témájának megválasztásánál.

## 2. Zajok és modellezésük

### 2.1. Zajok fajtái és forrásai

#### 2.1.1. Általános felosztás

A zaj definiálható olyan nem kívánt jelnek, mely megakadályozza (zavarja) egy másik jellel (hasznos jel) történő kommunikációt vagy a másik jel mérését. A zaj maga is egy információhordozó, a zajforrás szempontjából. Például egy autó motorjának zaja információval szolgál a motor állapotáról. A zajok igen sokfélék és változatosak lehetnek, kezdve a hangfrekvenciás akusztikus zajoktól a rádiófrekvenciás elektromágneses zajokig.

A másik említésre méltó probléma a hasznos jel kiindulási alakjától való nem kívánatos eltérése, torzulása. Ebbe a csoportba tartoznak a különféle visszhangok, felharmonikusok, és az elveszett jelerészetek.

A zajok és torzulások a kommunikációs és mérési rendszerek fő ellenségei, ezért e káros mellékhatások eltüntetése érdemel kiemelt figyelmet. A zajszűrést és torzítás kiküszöbölését főként a mobil kommunikációban, a beszéd felismerésben, orvosi jelfeldolgozásban, radar és hanglokátor rendszerekben alkalmazzák a köznap életben.

A zajok életünk minden területén, minden környezetben jelen vannak, különféle formában. Az zaj forrásától függően a következő nagy csoportok különíthetők el egymástól:

- a.) *akusztikus zaj*: leggyakrabban mozgásból, vibrációból (rezgésből), ütközésből származik, ezáltal ez a zajforma – eltérő mértékben ugyan – jelen van a mindennapi környezetünkben. Ezt a zajt legtöbbször haladó gépjárművek, légkondicionálók, számítógépek hűtőventillátorai, közlekedés, emberi beszélgetések, szél, eső, stb. okozzák.
- b.) *elektromágneses zaj*: minden frekvencián jelen van, különösen a rádiófrekvenciákon. Minden elektromos berendezés, például a rádió- és TV adóállomások, és vevőkészülékek állítanak elő jelentős elektromágneses zajt.
- c.) *elektrosztatikus zaj*: feszültség jelenléte okozza, áram nélkül vagy folyó árammal. Ennek a zajnak az egyik legjelentősebb forrása a fénycsővilágítás.

d.) *csatornatorzítás, visszhang, és csillapítás*: az átviteli csatorna nem ideális karakterisztikája következtében történik. A rádiófrekvenciás csatornák, melyeket a mobil kommunikációban használnak, – mikrohullámú frekvenciákon – különösen érzékenyek a terjedés karakterisztikájára és környezetére.

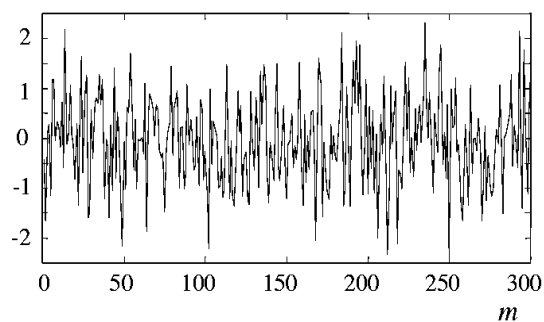
e.) *feldolgozási zaj*: az a zaj, amelyet a jel digitális–analóg átalakítása eredményez, például a zaj digitálisan kódolt beszédben vagy hasonló jelekben, vagy az elveszett csomagok a digitális adatkommunikációs rendszerekben.

A zaj frekvenciájának vagy időkarakterisztikájának függvényében a következő kategóriákba csoportosítható:

- *Keskeny hullámsávú zaj*: olyan keskeny hullámsávon érkező zajhatás, mint az 50 Hz-es vagy 60 Hz-es brummogás (búgás) mely többnyire az elektromos tápegységekből, váltakozó áramú vezetékekből származik.
- *Fehér zaj*: tisztán véletlen zaj, melynek teljesen egyenes az energiaspektruma. A fehér zaj elméletileg az összes frekvenciát tartalmazza, azonos intenzitással.
- *Sávkorlátozott fehér zaj*: egyenes spektrumú zaj korlátozott sáv szélességgel, mely általában az alkalmazott eszköz korlátozott szélességű sávját, vagy az általunk használt sávot jelenti.
- *Szinezett zaj*: nemfehér zaj, vagy egyéb szélessávú zaj, melynek nem egyenes a spektruma, például: rózsaszín zaj, barna zaj, autoregresszív zaj.
- *Impulzusszerű zaj*: véletlenszerű rövid időtartamú impulzusokból áll, véletlenszerű amplitúdóval, hosszal és felbukkanással.
- *Tranziens zaj*: viszonylag hosszú időtartamú impulzusokból áll.

### 2.1.2. Fehér zaj

A fehér zaj definíció szerint egy összefüggéstelen jelekből álló zaj, mely minden frekvencián azonos energiájú. A zaj, melynek energiája minden frekvencián azonos  $\pm \infty$  tartományban, szükségképpen végtelen nagy energiájúnak kellene lennie, ennek következtében ez csak egy elméleti fogalom. Habár egy sávkorlátozott zaj,



1. ábra Fehér zaj

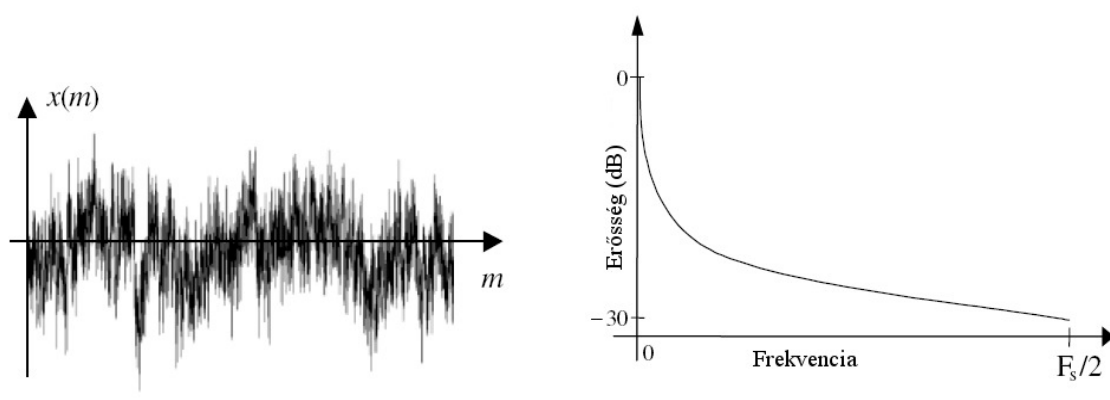
egyenes spektrummal – ha a zaj a teljes vizsgált sávot lefedi, amit a kommunikációs rendszer biztosít – a rendszer szempontjából minden tekintetben fehér zajnak számít. Például egy 10 kHz sávátvítellel rendelkező hangrendszer számára minden 10 kHz-nél szélesebb sávú egyenes spektrumú hallható zaj fehér zajnak

tűnik.

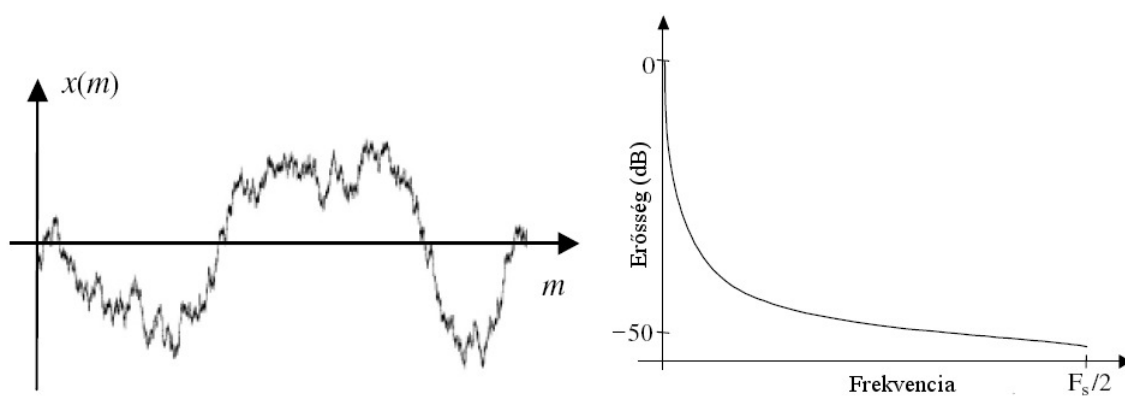
A tiszta fehér zaj egy elméleti fogalom, mivel végtelen energiájúnak kellene lennie a végtelen frekvenciatartomány lefedésére. Ráadásul a diszkrét időpontokban mintavételezett jelnek szükségképpen sávkorlátozottnak kell lennie, ahol a legnagyobb frekvenciaérték kisebb, mint a mintavételezési frekvencia fele. A sávkorlátozott fehér zaj egy sokkal gyakorlatibb fogalom, mely definíció szerint egy egyenes spektrumú zaj egy adott frekvenciasávban. Egy ilyen zaj energiája nem végtelen, ezért kiszámítható.

### 2.1.3. Színezett zaj

Bár a fehér zaj elmélete meglehetősen valós, és matematikailag kényelmes és használható közelítést ad több zajjal kapcsolatban, melyek felütik fejüket a telekommunikációs rendszerekben, sok egyéb zaj viszont nem fehér. A színezett zaj kifejezés bármely olyan szélessávú zajra utal, melynek spektruma nem a fehér zajéval egyező. Például a legtöbb hangfrekvenciás zaj, mint a mozgó autók zaja, a számítógépek ventilátorainak zaja, a fűrés zaja és emberi beszélgetés zaja, nemfehér, és túlnyomórészt alacsonyfrekvenciájú hangokat tartalmaz. Egy fehér zaj adott csatornán történő továbbításakor a zaj bizonyos frekvenciákon történő csillapodásával jellemezhető a „színezett zaj”, a csatorna spektrumának alakjától függően. A színezett zaj két klasszikus változata a rózsaszín (2. ábra) és a barna (3. ábra) zaj.



2. ábra Rózsaszín zaj és spektruma



3. ábra Barna zaj és spektruma

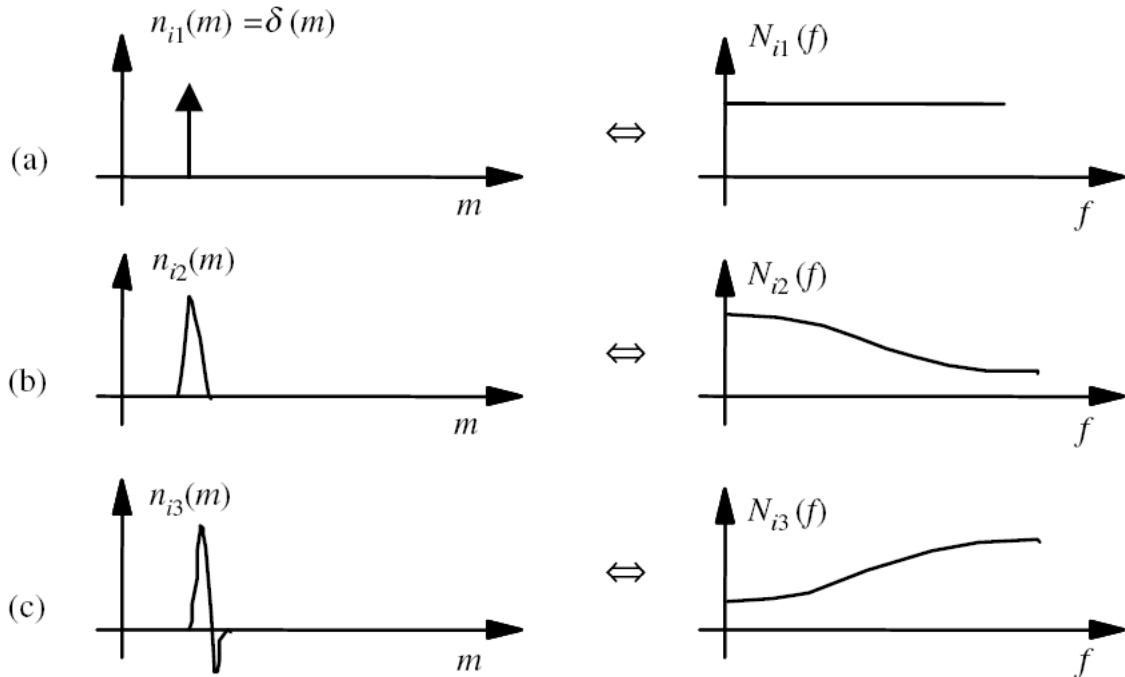
#### 2.1.4. Impulzív zaj

Az impulzív zaj rövid időtartamú „be/ki” impulzusokból áll. Ezek a zajok többféle forrásból származhatnak, mint például kapcsolási zaj, kedvezőtlen csatorna-környezet, kimaradás vagy felületi kopás a hanghordozón, számítógép-billentyűzet kattogása, és így tovább. A 4. ábra (a) része egy ideális impulzív zajt mutat, annak spektrumával. A kommunikációs rendszerekben előforduló impulzív zaj jellemző paramétere a hossza, mely normális esetben több mint egy minta hosszúságú. Például hangrendszerek esetében a rövid időtartamú, éles impulzusok lehetnek akár 3 milliszekundum (60 minta 20 kHz-es mintavételezési frekvencia esetében) hosszúak, melyeket impulzív zajnak tekinthetünk. A 4. ábra (b) és (c) része két példát mutat rövid időtartamú impulzusokra és azok spektrumára.

Egy kommunikációs rendszerben az impulzív zaj a térben és időben több pontból ered, azután a csatornán keresztül halad a vevő felé. A vett zaj időben szétszórt és a csatorna által alakított, és úgy kezelhető, mint a csatorna impulzus jelre adott reakciója. Általában a csatornák karakterisztikája lehet lineáris vagy nem lineáris,

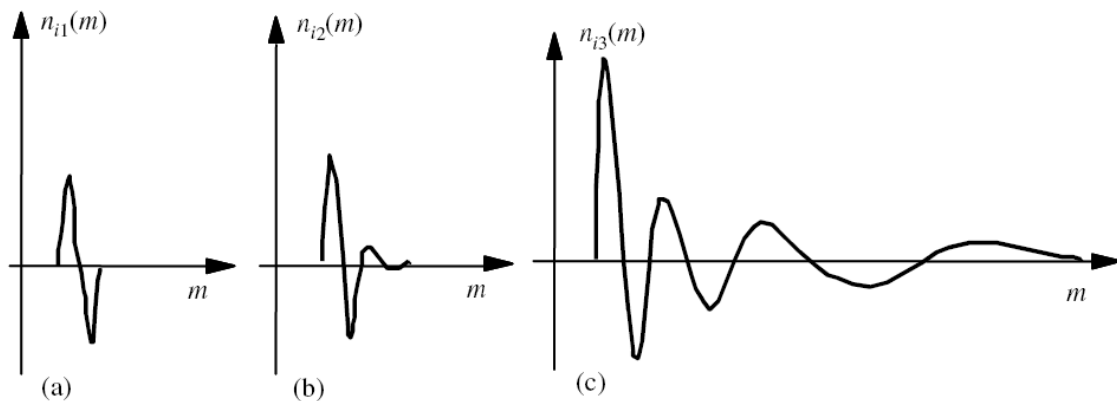


időben állandó, vagy időben változó. Azonkívül sok kommunikációs rendszer válaszul a nagy amplitúdójú impulzusokra nem lineáris karakterisztikát mutat.



4. ábra Impulzív zaj

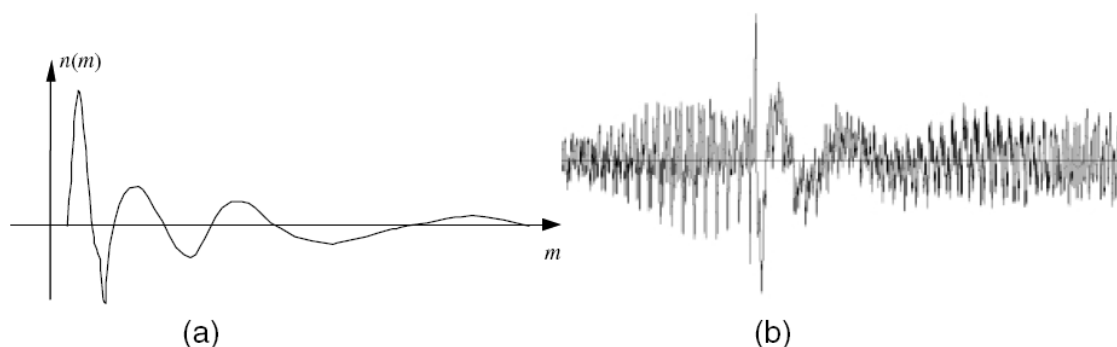
Az 5. ábra néhány példát mutat impulzív zajra, melyek tipikusan régi hanglemezeken fordulnak elő. Ebben az esetben a kommunikációs csatornának a lejátszó rendszert tekintjük, melyet időben állandónak feltételezhetünk. Az ábra három változatát mutatja a zaj-amplitúdóváltozásának az idő függvényében. Az 5. ábra (c) eleme mutatja, hogy egy nagy impulzslökés csillapodó tranziens impulzussorozatot hoz létre. Ezek a példák a lejátszó rendszer nem lineáris karakterisztikáját jelzik.



5. ábra Példák impulzív zajra

### 2.1.5. Tranziens zaj impulzus

Az tranziens zaj gyakran áll egy viszonylag rövid éles impulzusból, melyet egy csillapodó viszonylag alacsony frekvenciájú oszcilláció követ, ahogy azt a 6. ábrán is láthatjuk. A kezdő impulzust gyakran valamilyen belső vagy külső interferencia okozza, míg az oszcillációt gyakran a csatorna rezonanciája, melyet a kezdeti impulzus gerjeszt, és tekinthető a csatorna válaszának a kezdeti impulzusra. Egy telekommunikációs rendszerben az impulzív zaj a térben és időben több pontból származik, azután a csatornán keresztül halad a vevő felé. A vett zaj időben szétszórt és a csatorna által torzított, alakított és úgy tekintendő, mint a csatorna egy impulzusra adott válaszreakciója. Az impulzív zaj a kezdő impulzussal és az azt követő oszcillációval jellemezhető, a csatorna szempontjából.



6. ábra Egy karc impulzus és zene egy hanglemezről (b), hanglemez karc impulzusának átlagolt profilja (a)

A 6. ábra jól szemlélteti a tranziens zaj impulzus alakját, tekintsük meg a karcolás impulzust<sup>1</sup> az ábrán, mely egy sérült hanglemezről származik. A karcolás impulzusok a lemeztű és a hozzá kapcsolt elektromechanikus lejátszó rendszer akusztikus válasza a lemezen található fizikai megszakítások jelenlétére. Mivel a karcolás hatására előálló tranziens zajok tulajdonképpen a lejátszó rendszer impulzus válasza, elvárható, hogy az adott rendszer különböző helyeken található impulzusokra hasonló karakterisztikát mutasson. Ahogyan a 6. (a) ábra is mutatja egy tipikus karc impulzus hullámalak gyakran két különböző részre tagolódik:

- egy bevezető nagy amplitúdójú lökésre, mely a lejátszó rendszer válasza a rögzített médiumon (lemezen) található karcra, melyet
- csillapodó hullámzások követnek, további torzítást okozva ezzel.

<sup>1</sup> Karc(olás) impulzus alatt egy hanglemezen, a használat során kialakult (többnyire a sávval valamilyen szöveget bezáró) karc (tekinthető pillanatnyi jelfolytonosság-hiánynak is) hatására létrejövő lejátszáskor jelentkező ugrásszerű lökést érthetünk, mely közvetlenül a mechanikus jelérzékelőre hat, és a mechanikai–elektronikai átalakítást követően végighalad a teljes elektronikus rendszeren is.

A kezdő impulzus általában viszonylag rövid, és a hossza a 1–5 ms tartományba esik, míg a hullámzó farok hosszabb időtartama elérheti akár az 50 ms-ot vagy többet is.

A 6. (b) ábrán látható csillapodó hullám (farok) frekvenciája is csökken az idő múlásával. Ez a viselkedés az elektromechanikus lejátszó-rendszer nem lineáris válaszmechanizmusa a lemezen található karcokra. Tapasztalat alapján több hanglemezről származó recsenés zaja jól meghatározható hullámalakot mutat, és viszonylag kevés tipikus sablonnal jellemezhető.

### 2.1.6. Termikus zaj

A termikus zaj – más néven Johnson-zaj (felfedezőjéről J. B. Johnsonról<sup>1</sup> elnevezve) – hővel stimulált részecskék rendezetlen mozgásából adódik. A termikus zaj elméletének alapjai a termodinamikában keresendők. A termodinamikában a szabad részecskéket hőmérsékletfüggő rendezetlen mozgás jellemzi, mint például a szabad gázmolekulák egy tartályban, vagy a szabad elektronok egy vezetőben. Bár ezen rendezetlen mozgások átlagértéke közel nulla, az átlag körüli hullámzások alkotják a termikus zajt. Például a gázmolekulák rendezetlen mozgásai és ütközései egy behatárolt térben rendezetlen hullámzást okoznak az átlagos nyomás körül. A hőmérséklet növekedésével a molekulák vagy más elemi részecskék kinetikai energiája növekszik, és ezzel a termikus zaj is.

Ehhez hasonlóan egy elektromos vezetőben nagy számú szabad elektron van jelen (az ionokkal együtt), melyek rendezetlenül rezegnek az egyensúlyi pozíciójuk körül, és útját állják a további elektronok mozgásának. Az elektronok rendezetlen szabad mozgásai spontán áramokat, vagy más szóval termikus zajt képeznek, melynek átlaga nulla, mivel feszültség hiányában minden elektron mozgása különböző irányú (nincs egyirányú áramlás). A környezet hőmérsékletének növekedésével a vezető hőmérséklete is növekszik, mely az elektronokat magasabb energiaszintre viszi, ezzel növelve a rendezetlen áramfolyásokat. Egy fémréteg ellenállásban a pillanatnyi feszültségek négyzetének átlagértéke a termikus zaj miatt a következő összefüggéssel adható meg:  $\overline{v^2} = 4kTRB$ , ahol  $k = 1,38 \cdot 10^{-23} \frac{\text{J}}{\text{°K}}$  a Boltzmann-állandó,  $T$  az abszolút hőmérséklet °K-ben,  $R$  az ellenállás ohmban, és  $B$  a sáv szélesség. Az iménti egyenletből és az azt megelőző érvelésből egy áramkörü

<sup>1</sup> John Bertrand “Bert” Johnson (1887–1970): svéd születésű amerikai villamosmérnök és fizikus. Ő fejtette ki először részletesen a kábeleken közvetített információ véletlen interferenciájának alapjait. (Wikipédia – [http://en.wikipedia.org/wiki/John\\_B.\\_Johnson](http://en.wikipedia.org/wiki/John_B._Johnson))

lapba ültetett fémréteg ellenállás úgy viselkedik, mint egy termikus zajgenerátor, mely feszültségnégyzetének átlaga  $\overline{v^2}$ , és a belső ellenállása  $R$ . Az ily módon előálló termikus zajgenerátor teljesítménye számítható ki a következő összefüggéssel:

$$P_N = \overline{i^2} R = \left( \frac{v_{rms}}{2R} \right)^2 R = \frac{\overline{v^2}}{4R} = kTB [\text{W}], \text{ ahol } v_{rms}^1 \text{ az effektív feszültség értéke. A ter-}$$

mikus zaj spektrumsűrűsége megadható a  $P_N(f) = \frac{kT}{2} \left[ \frac{\text{W}}{\text{Hz}} \right]$  összefüggéssel. Ebből

az egyenletből kiderül, hogy a termikus zaj spektrumsűrűsége egyenes, azaz a termikus zaj egy fehér zajnak tekinthető (mely általában sávkorlátozott). Ezt az összefüggést különösen a magas rádiófrekvenciák ( $10^{13}$  Hz) esetében fontos.

### 2.1.7. Schottky-zaj

A Schottky-zaj<sup>2</sup> elmélete az elektroncsövek katódjának elektronkibocsátás-ingadozásából származik. Az különállónak tekintett elektronok áramlás során véletlenszerű időpillanatokban érkeznek, ezért az átlagérték körüli ingadozást okoznak a részecskefolyamban. Ezek az ingadozások a részecskefolyam sebességében Schottky-zajt okoznak. Egyéb példák Schottky-zajra a fotonok áramlása egy lézer fénysugárban, elektronok és lyukak áramlása valamint rekombinációja<sup>3</sup> félvezetőben, vagy a fotóelektronok árama, melyek a fotódiódákban emittálódnak. A véletlenszerűség jelenléte a részecske-kibocsátás vagy -érkezés mértékében magába foglalja, hogy a Schottky-zaj Poisson eloszlással modellezhető. A hosszú megfigyelési idejű kísérletek esetén az érkezések átlagérték körüli ingadozása a normális eloszláshoz közelít. Míg a termikus zaj a részecskék „nyugalmi állapotban” történő véletlenszerű mozgásából adódik, addig a Schottky-zaj egy (feszültséggel) irányított részecske-áramlásra értelmezhető.

<sup>1</sup>  $v_{rms}$  – root mean squared voltage: a feszültség effektív értéke

<sup>2</sup> Walter Hermann Schottky német fizikus, a film-rácsú vákuumcső (1915) és a tetróda (1919) feltalálója – munkáit a Siemensnél végezte. 1938-ban megfogalmazott egy elméletet, mely a Schottky-hatást jelezte, ma pedig a Schottky-diódákban használjuk. – [http://en.wikipedia.org/wiki/Walter\\_H.\\_Schottky](http://en.wikipedia.org/wiki/Walter_H._Schottky)

<sup>3</sup> „A félvezető nem csak szabad elektronjai útján vezeti az áramot (mint a fémek), hanem egy másik vezetési mechanizmus: a lyukvezetés útján is. Az eredetileg elektromosan semleges atom a negatív töltésű elektron távozása folytán pozitív töltésűvé válik, ezért a lyukat pozitív töltésűnek tekintjük. Előfordulhat, hogy egy szomszédos atom valamelyik vegyértékelektronja szintén elszabadul, és „beugrik” az előbb említett (pozitív töltésénél fogva az elektront vonzó) lyukba. Ezzel az előbbi lyuk (és az oda beugró szabad elektron is) megszűnik, hiszen a rácsszerkezet rendje az adott helyen helyreállt (ezt a folyamatot nevezik *rekombinációnak*), viszont a most elszabadult szabad elektron helyén új lyuk keletkezett. Ha valamelyik szomszédos atomból elszabaduló elektron most ebbe a lyukba ugrik be, tehát rekombinálódik, a lyuk a most elszabadult atom helyére »vándorol át«.” Forrás: Internet – [http://www.puskas.hu/r\\_tanfolyam/felvezetok.pdf](http://www.puskas.hu/r_tanfolyam/felvezetok.pdf), 2. oldal

Tekintsünk egy elektromos áramot, mint különálló elektromos töltéshordozók áramlását. Ha a töltéshordozók egymástól függetlenül mozognak, akkor az áram ingadozása megadható a  $I_{Noise}(rms) = \sqrt{2eI_{dc}B}$  összefüggéssel, ahol  $e = 1,6 \cdot 10^{-19}$  coulomb (egy elektron töltése), és  $B$  pedig a mérhető sávszélesség. Például egy  $I_{dc}$  1 amper „állandó” áram, és 1 MHz sávszélesség mellett, az effektív hullámzás 0,57 mikroamper lesz. Az iménti összefüggés feltételezi, hogy egymástól független töltéshordozók alkotják az áramot. Abban az eset mikor a töltéshordozóknak egy akadályon kell áthaladniuk, mint például egy félvezető dióda p–n átmenete, a töltéshordozók nem együtt mozognak; viszont nem igaz ez a fém vezetőkre, ahol a töltéshordozókat nagyfokú kölcsönhatás jellemzi.

### 2.1.8. Elektromágneses zaj

Virtuálisan minden elektromos eszköz, mely energiát állít elő, energiát fogyaszt vagy energiát szállít, elektromágneses zajok potenciális forrása. Az ilyen elektromágneses zaj interferenciának tekinthető más rendszerek szempontjából. Általában minél nagyobb a feszültség vagy az áram és minél közelebb helyezkedik el az áramkörhöz vagy elektromos eszközhöz, annál nagyobb lesz az áramkörben tapasztalható elektromágneses zaj mértéke. Az elektromágneses zajok általános forrásai a transzformátorok, rádió és televízió adóállomások, mobiltelefonok, mikrohullámú adóállomások, váltakozó áramú vezetékek, motorok és motorindítók, generátorok, relék, oszcillátorok, fénycsövek, és villámlások (zivatarok).

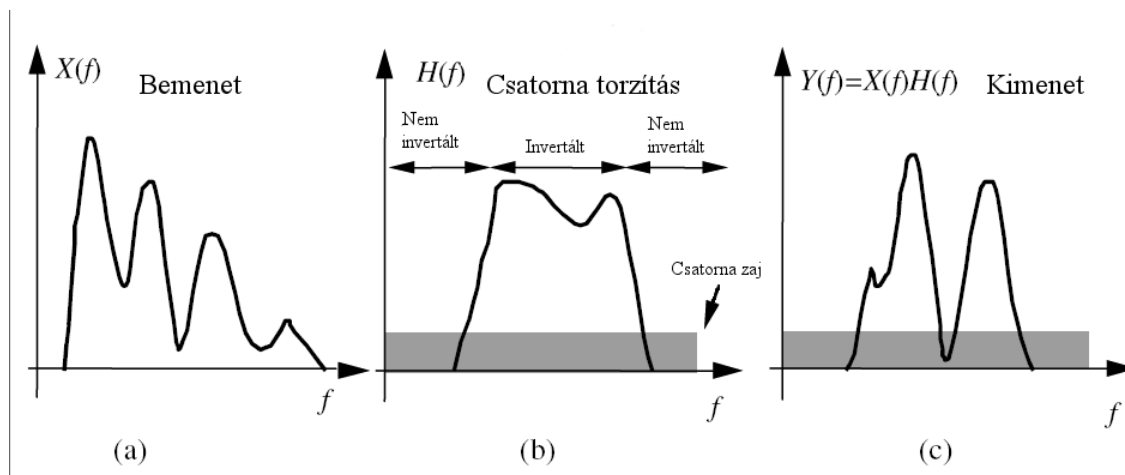
Az elektromágneses zajok két alapvető csoportba sorolhatók: elektrosztatikus zajok, és mágneses zajok. Ez két típus alapvetően különbözik egymástól, ezért más-más zajvédelem szükséges hozzájuk. Sajnos az imént felsorolt zajforrások legtöbbször a két típus kombinációját állítja elő, mely tovább bonyolítja a zajcsökkentési problémát.

Az elektrosztatikus tereket a feszültség jelenléte generálja, árammal, vagy anélkül. Az elektrosztatikus zajok egyik legáltalánosabb forrásai a fénycsövek. Mágneses terek keletkeznek elektromos áram hatására, vagy állandó mágnes jelenléte esetén. A motorok és a transzformátorok szolgálhatnak például az iménti felsorolásból, és hogy ne menjünk messzire a Föld mágneses tere is itt van környezetünkben. Ahhoz, hogy zajfeszültség indukálódjon egy vezetőben, a mágneses erővonalakat metszenie kell. A generátorok működési elve is ezen a jelenségen alapul. A váltakozó mágneses tér jelenléte – amilyen például az 50/60 Hz-es elektromos vezetékeket körülveszi – feszültséget indukál minden álló vezetőben, ahogy a mág-

neses tér kialakul, majd eltűnik (váltakozó mágneses tér). Ehhez hasonlóan, ha egy vezető mozog a Föld mágneses terében, zajfeszültség indukálódik benne, ahogy mozgás közben metszi a Föld mágneses erővonalait.

### 2.1.9. Csatornatorzítások

Ha egy jelet egy csatornán keresztül továbbítunk, a jel alakja változik, torzul a csatorna frekvenciaátvivő képességének és csillapítási paramétereinek megfelelően. A csatornatorzítás két fő megnyilvánulási formája a folytonossági torzítás és a fázistorzítás. A rádiókommunikációban létezik egy multi-út hatás is, mely szerint ha a jel több átjátszón megy keresztül, bizonyos esetekben különböző utakon haladva, ezzel a jel többféle változata érkezik a vevőbe különböző késleltetéssel és csillapítással. A csatornatorzítások csökkenthetők, vagy egészen egyszerűen szétválaszthatók a kommunikációs folyamatokban, ezért a csatorna modellezése és ki-

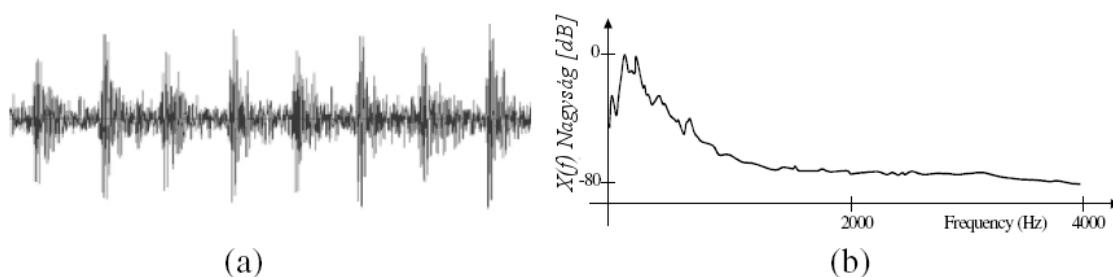


7. ábra Csatornatorzítás

egyenlítése nagyon jelentős összetevői egy modern digitális kommunikációs rendszer kialakításának. A csatorna kiegyenlítése különösen fontos a modern sejtes felépítésű kommunikációs rendszerekben mivel a csatorna karakterisztikák és terjedési paramétereik különbsége igen változatos, sokkal változatosabb, mint a vezetékes rendszerekben. A 7. ábra egy csatorna sávátviteli jellemzőit mutatja, mely két nem megfordítható és egy megfordítható régióra oszlik. A megfordítható csatornarégióban a jelek erősen torzultak elvesztek a csatornazaj miatt. A nem megfordítható csatornarégióban a jelek torzultak ugyan, visszanyerhetők. Ez a példa remekül mutatja, hogy a csatornafordító szűrőket nagy gondossággal kell elkészíteni, a nem kívánt eredmények elkerülése érdekében, mint például a zajerősítés az alacsony jel-zaj arányú frekvenciákon.

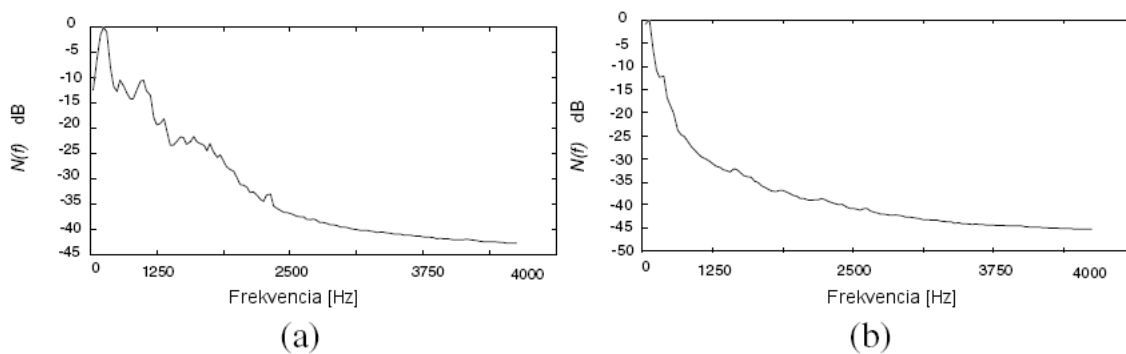
## 2.1.10. Zajmodellezés

A cél a modellezés esetén, hogy a jelben vagy zajban jelalakokat és mintákat jellemezzünk. A zaj pontos modellezéséhez egy olyan struktúrára van szükségünk, mely mind időben mind spektrumában jellemzi a zajt. A pontos zajstatisztika modellezéshez a kulcs a zajos jel megfelelő osztályozása és a hasznos jel erősítése. Ezen látszólag egyszerű jel/zaj osztályozási munka döntően függ a rendelkezésre álló hasznos jel mennyiségétől, erősségétől és a zajmodellektől, és ezen modellek alkalmazásától.



9. ábra A fúrózaj (a) idő-hullámformája (b) frekvenciaspektruma

A leghasználhatóbb szinte nélkülözhetetlen módszer az összes közül a zaj felépítésében való éleslátásra a Fourier transzformáció a zaj frekvenciaanalizálására. A 8. ábra egy elektromos fúró zaját mutatja, és ahogy az várható a zaj periodikus. Az ábra (b) része mutatja a fúró zajának spektrumát, melyen látható, hogy a zaj energiájának jelentős része a spektrum alacsonyabb frekvenciáira koncentrálódik. Tényszerűen igaz, hogy a legtöbb hangfrekvenciás jelnek és zajnak van túlnyomórészt alacsonyfrekvenciás a spektruma. Bár meg kell jegyezni, hogy a viszonylag kisebb energiájú magasabb frekvenciájú hangjelek fontos szerepet játszanak az élmény és a minőség átadásában. A 9. ábra (a) és (b) része példával szolgál az autók belsejéből származó zaj spektrumából (egy BMW és egy Volvo zajai láthatók). A



8. ábra Autók zajának spektuma (a) BMW 112 km/h, (b) Volvo 112 km/h

zajok nem statikusak, változatosak, és a következő forrásokból származhatnak:

- a.) kvázi periodikus zaj a motorból, és a jármű forgó alkatrészeiből,
- b.) zaj a kerekek és a talaj érintkezéséből és az út felületéből adódóan,
- c.) az autó körüli légáramlatok a légjáratokon, ablakokon, napfénytetőn keresztülhaladva,
- d.) az éppen előzőtt jármű zaja

Az autó zajainak jellemzői változnak a sebességgel, az út felületének állapotváltozásával, az időjárással és az autó környezetével.

A zajmodellezés legegyszerűbb módja, melyet gyakran használunk a jelenlegi gyakorlatban, hogy kiszámítjuk a zaj statisztikáját jelmentes időszakok alapján. Optimális esetben a Bayes féle jelfeldolgozási módszerekben egy készlet valószínűségi modell van előkészítve a jel- és zajfeldolgozásra. Ezek a modellek alapjául szolgálnak a jelek és zajok dekódolásának, zajos jelek felismerésének, erősítésének.

#### 2.1.10.1. Hozzáadódó fehér normális eloszlású zajmodell (AWGN<sup>1</sup>)

A kommunikációelméletben a gyakran feltételezzük, hogy a zaj hozzáadódó fehér normális eloszlást követő. Bár számos probléma esetében ez valós feltevés, és matematikailag kényelmes és használható megoldáshoz vezet, a gyakorlatban a zaj gyakran időben változó, kölcsönösen összefüggő, és nem normális eloszlást követő. Ez részben igaz az impulzív típusú zajokra, melyek nem állandóak, nem normális eloszlásúak, ezért nem modellezhetők az AWGN feltevés használatával. A nem állandó és nem normális eloszlású zajok modellezésére a Markov-féle sorozat alkalmas, stacionárius alfolyamatok leírásával.

#### 2.1.10.2. Rejtett Markov modell

A legtöbb zaj nem állandó, ez érvényes a zaj statisztikai paramétereire, úgy mint a zaj középértéke, szórásnégyzete, erőssége és spektruma, és ezek időbeli változása. A nem stacionárius zajok rejtett Markov modellek (HMM<sup>2</sup>) segítségével modellezhetők. Egy HMM lényegében stacionárius alfolyamatok véges sok állapotú Markov sorozata. Az állandó zaj modellezhető egy egyállapotú HMM segítségével. A nem állandó zaj modellezésére a többállapotú HMM alkalmas, amely képes a zaj időbeli változásának követésére, véges sok állandó állapottal. Nem normális eloszlású zajok esetében egy Gauss sűrűségi modell keverékét alkalmazhatjuk, a zaj terének modellezésére minden állapotban. Általában az állapotok száma modellen-

---

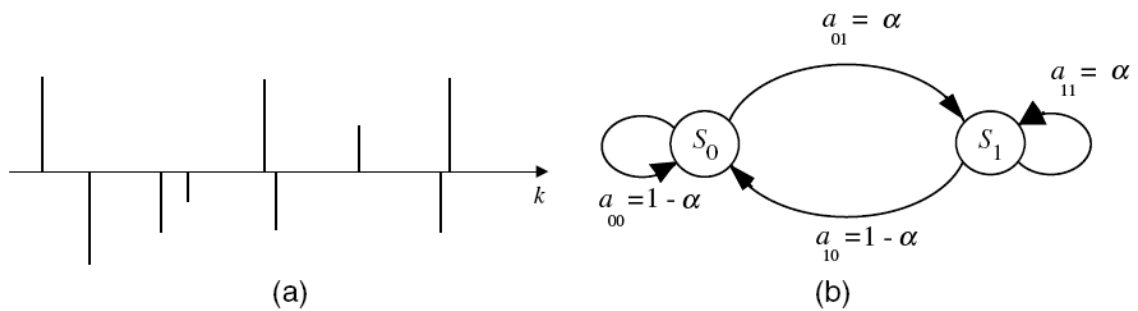
<sup>1</sup> AWGN: Additive White Gaussian Noise Model

<sup>2</sup> HMM: Hidden Markov Model



ként, és a keverékek száma állapotonként szükséges a zaj pontos modellezéséhez, mely a nem állandó zaj karakterisztikájától függ.

Egy példa a nem állandó zajra az impulzív zaj, melyet a 10. ábra (a) része szemléltet. Az ábra (b) részén az impulzív zaj szekvencia egy kétállapotú HMM modellje látható, ahol az  $S_0$  állapot jelképezi az „impulzus-ki” periódusokat az impulzusok között, és az  $S_1$  állapot jelképezi magát az impulzust. Azokban az esetekben ahol minden impulzus jól definiált átmeneti felépítéssel rendelkezik, előnyös lehet a többállapotú HMM magának az impulzusnak a modellezésére.



10. ábra (a) Impulzív zajszekvencia, (b) az impulzív zaj kétállapotú modellje

## 2.2. Zajmodellek

### 2.2.1. Valószínűségi modellek

A valószínűségi modellek alapja az információelmélet. Az információ maga is mennyiségileg a valószínűség logaritmusaként fejezhető ki. A valószínűségi modellek általában azzal foglalkoznak, hogy történések előfordulását megjósolják és jellemzésük véletlenszerű események alapján, különféle területeken, mint például telefonhívások számának megjósolása egy fővonalon a nap egy meghatározott időszakában, közúti forgalommodellezés, időjárás előrejelzés, pénzügyi adatmodellezés, előre jelezni a drogok hatását megadott adatok alapján az orvosi gyakorlatban. A jelfeldolgozásban a valószínűségi modelleket általában véletlenszerű jelek ingadozásának leírására alkalmazhatjuk, mint például mintafelismerés, jelkódolás és jelkiértékelés.

#### 2.2.1.1. Véletlenszerű jelek és sztochasztikus folyamatok

A jelek egyik legalapvetőbb jellemzőjük szempontjából két nagy csoportba osztályozhatók: determinisztikus (pontosan meghatározott) jelek és véletlenszerű jelek. Az időben véletlenszerű jelekre gyakran úgy hivatkozunk, hogy sztochasztikus jelek. Minden osztályban a jelek lehetnek folyamatosak és diszkrét az időben, és rendelkezhetnek folytonos értékű és diszkrét értékű amplitúdóval.

A determinisztikus jel definiálható úgy, hogy egyike az előre meghatározott görbéjű jeleknek időben és térben. A determinisztikus jelek pontos hullámszáma teljes mértékben leírható az idő függvényében, és a jel pontos értéke bármely időpillanatra kiszámítható a függvényi leírásból, és a korábbi értékek is kiszámíthatók ezen a módon. Például egy szinusz hullám  $x(t)$  modellezhető és pontosan kiszámítható mind egy másodrendű lineáris prediktív modellből, vagy a közismertebb  $x(t) = A \cdot \sin(2\pi ft + \phi)$  egyenlettel.

A véletlenszerű jeleknek nem kiszámítható hullámszámai vannak, ezért nem lehet olyan egyenletet felírni, mellyel kiszámítható lenne a jel pontos jövőbeli értéke a korábbi értékeiből. A legtöbb jel, mint például a beszéd, vagy zaj legalább részben véletlenszerű. A véletlenszerűség elmélete az információ és a zaj elméletével teljesen összekapcsolódott. Valóban, a véletlenszerű jelek feldolgozásakor a munka nagy részét jelenti az információ kinyerése a zajosan vett jelből. Ha egy jel alkalmas információhordozásra, akkor valamilyen fokú véletlenszerűségnek kell benne lennie: a kiszámítható jelek nem hordoznak információt. Ezért a jel véletlenszerű

része lehet a jel információtartalma, vagy zaj, vagy a kettő keveréke, tehát információ és zaj. Bár a véletlenszerű jelek nem teljesen kiszámíthatók, mégis gyakran mutatnak egy sor olyan statisztikailag jellemző értéket, mint például a maximum, a minimum, a középérték, a felező, a szórásnégyzet, és a teljesítményspektrum. A véletlenszerű jel leírható a statisztikája szempontjából, és a legteljesebben a valószínűségi modell szempontjából, melyből az összes statisztikája számítható.

#### 2.2.1.1.1. Sztochasztikus folyamatok

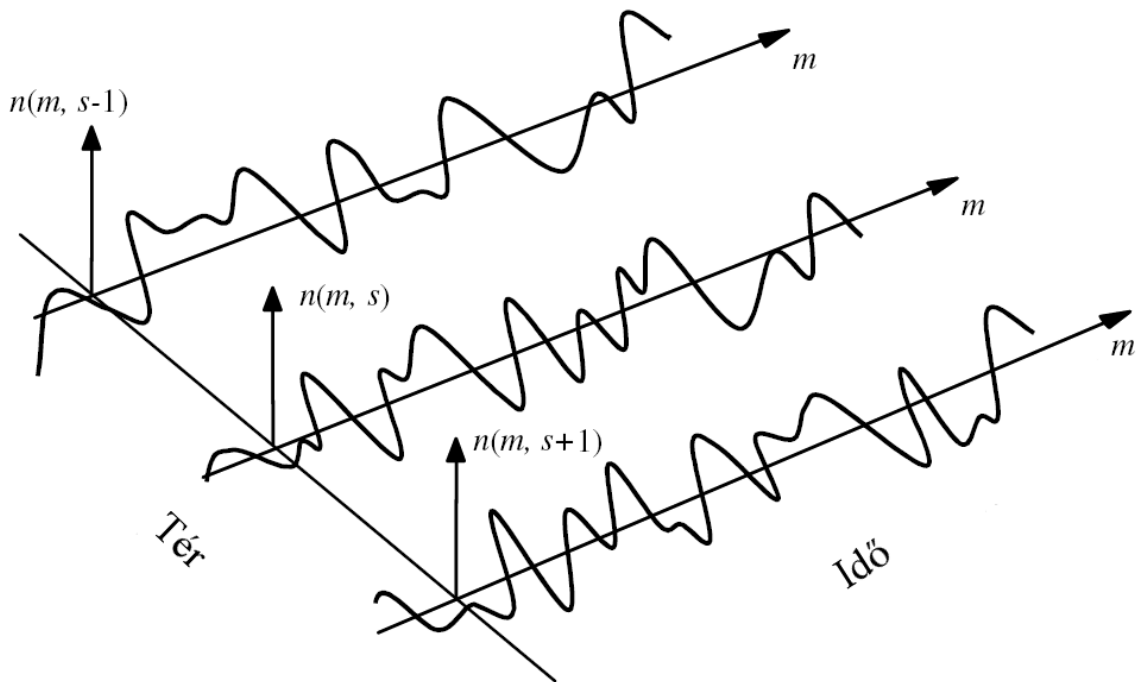
A „sztochasztikus eljárás” szakkifejezést általában a véletlenszerű folyamatok leírására használjuk, melyek folyamatos jeleket állítanak elő, mint például az emberi beszéd, vagy zaj. A jelfeldolgozási terminológiában egy sztochasztikus eljárás egy valószínűségi modell, a véletlenszerű jelek egy osztályára, például Gauss módszer, Markov módszer, Poisson módszer. Egy klasszikus példa a sztochasztikus eljárásra a részecskék úgynevezett Brown mozgása folyadékokban. A részecskék a folyadék belsejében véletlenszerűen mozognak a folyadékrészecskék „bombázása” következtében. Minden részecske véletlenszerű mozgása a sztochasztikus módszer egy egyszerű elképzelése. Minden részecske folyadékban történő mozgása segíti a megértést, vagy a módszer különféle megvalósításainak terét.

#### 2.2.1.1.2. Véletlen folyamatok együttese vagy tere

A véletlenszerű folyamatok összes elképzelésének gyűjteménye a módszerek együttese vagy tere néven ismert. Példaként tekintsünk egy véletlenszerű zajt egy telekommunikációs rendszerben ahogyan azt a 11. ábra is mutatja. A zaj minden telefonvonalon időben véletlenszerűen ingadozik, és jelölhető az  $n(m,s)$  jelöléssel, ahol  $m$  a diszkrét idő mérőszám, és  $s$  jelöli a sorindexet. A zajok gyűjteménye a különböző vonalakon a zaj együttesét (vagy terét) alkotják, melyet a  $N(m)=\{n(m,s)\}$  jelöléssel jelölünk, ahol  $n(m,s)$  az  $N(m)$  zaj realizációját jelöli az  $s$  vonalon.

#### 2.2.1.2. Valószínűségi modellek

A valószínűségi modellek a véletlenszerű jelek legteljesebb matematikai leírását adják. Egy  $m$  rögzített időpillanatban egy véletlenszerű jeltől gyűjtött minta  $\{x(m,s)\}$  egy valószínűségi változó, mely különböző értékeket vesz fel az  $s$  térben. Az alapvető különbség a véletlenszerű változó és a véletlenszerű folyamat között, hogy az utóbbi idősorozatot generál. Ezért a véletlenszerű változók modellezésére használt modellek használhatók a véletlenszerű folyamatokra is.



11. ábra A véletlen zaj három realizációja a zajok terében

A valószínűségi változó tere az összes érték gyűjteménye, vagy eredmény, amit a változó felvehet. A valószínűségi változó tere részekre osztható, néhány kritérium szerint, néhány altérre. Az altér jelértékek egy gyűjteménye, melyek egy közös tulajdonsággal rendelkeznek, mint egy csoportba összezsúfolt minta, vagy a minták gyűjteménye azok amplitúdójával az adott sávon belül. Minden alteret egy eseménynek hívunk, és az esemény bekövetkezésének valószínűsége  $A$ , és  $P(A)$  az a hányados, mely az  $A$  térből származó megfigyelések eredményeinek számát  $N_A$ -t

elosztjuk az összes megfigyelés számával:  $P(A) = \frac{N_A}{\sum_i N_i}$ . Az iménti egyenletből evidens, hogy a kísérletben résztvevő összes lehetséges esemény valószínűségének összege egy.

#### 2.2.1.2.1. Valószínűségi változó eloszlása, eloszlásfüggvény

Legyen adott egy  $(\Omega, \mathbf{A}, P)$  valószínűségi mező, legyen értelmezve egy  $\xi: \Omega \rightarrow \mathbf{R}$  függvény, melyre teljesül, hogy  $\{\omega \mid \xi(\omega) = x\} \in \mathbf{A}$  minden  $x \in \mathbf{R}$  esetén. ( $\xi$  az elemi események halmazát képezi le  $\mathbf{R}$ -be, értékkészlete diszkrét értékekből áll). Ekkor  $\xi$ -t diszkrét valószínűségi változónak nevezzük.  $\xi$  értékkészlete  $\{x_1, x_2, \dots\}$  megszámlálható.  $\xi$  eloszlása:  $\{p_i = P(\xi = x_i) = P(\omega \mid \xi(\omega) = x_i) \mid i=1, 2, \dots\}$  valószínűségek halmaza, annyi elemű, ahány elemű az értékkészlet, elemei mutatják a diszkrét valószínűségi változó egyes értékeinek felvételi valószínűségét. Az értékkészlet elemei nemnegatív számok, melyek összege 1.

Két diszkrét valószínűségi változó ( $\xi$  és  $\eta$ ) *együttes eloszlásán* a következő eloszlást értjük (amennyiben  $\xi$  értékészlete  $\{x_1, x_2, \dots\}$  eloszlása  $\{P(\xi=x_k) \ (k=1, 2, \dots)\}$ ,  $\eta$  értékészlete  $\{y_1, y_2, \dots\}$  eloszlása  $\{P(\eta=y_j) \ (j=1, 2, \dots)\}$ ):  $\{P(\xi=x_k \text{ és } \eta=y_j) \ (k,j=1, 2, \dots)\}$ .

Az együttes eloszlás elemei nemnegatívak, az elemek összege 1. Az együttes eloszlásban szereplő események teljes eseményrendszer alkotnak, ezért az előző állítás igaz.

A peremeloszlás két diszkrét valószínűségi változó együttes eloszlásából meghatározható az alkotó diszkrét valószínűségi változók eloszlása. Valamelyik diszkrét valószínűségi változó eloszlása szerint összegezve az együttes eloszlást, a másik diszkrét valószínűségi változó eloszlásához jutunk.

Diszkrét valószínűségi *változók függetlenek*, ha minden  $j = 1, 2, \dots$  és minden  $k = 1, 2, \dots$  esetén  $P(\xi=x_k \text{ és } \eta=y_j)=P(\xi=x_k)P(\eta=y_j)$ , azaz együttes eloszlásuk minden eleme szétbomlik az eloszlások megfelelő elemeinek szorzatára. Két diszkrét valószínűségi változó eloszlása pontosan akkor határozza meg együttes eloszlásukat, ha függetlenek. A konstans értékű valószínűségi változó minden más diszkrét valószínűségi változótól független.

Adott valószínűségi mezőben lévő  $B_i \ (i=1, 2, \dots)$  teljes eseményrendszer esetén ha ezek között nincs 0 valószínűségű esemény, akkor tetszőleges nem 0 valószínűségű  $A$  eseményre  $P(A | B) = \frac{P(A | B_i) \cdot P(B_i)}{\sum_i (P(A | B_i) \cdot P(B_i))}$ . Ez Bayes tétele.

Egy  $\xi$  valószínűségi változó *eloszlásfüggvénye*:  $F: \mathbf{R} \rightarrow [0, 1]$ ,  $F(x)=P(\xi < x)$ . Valószínűségi változó eloszlásfüggvénye balról folytonos, monoton nemcsökkenő,  $-\infty$ -ben 0-hoz tart,  $+\infty$ -ben 1-hez tart. Bármely, a valószínűségi változó eloszlásfüggvényeire vonatkozó fenti feltételeket kielégítő függvényhez létezik olyan valószínűségi változó, melynek eloszlásfüggvénye az adott függvény. A diszkrét valószínűségi változónak is létezik eloszlásfüggvénye, ez „lépcsős” alakú.  $\xi$  valószínűségi változó esetén  $P(a \leq \xi < b)=P(\xi < b)-P(\xi < a)=F(b)-F(a)$ , ha  $a < b$ .  $\xi$  és  $\eta$  *valószínűségi változók együttes eloszlásfüggvénye*  $F(x,y)=P(\xi < x \text{ és } \eta < y) \ (x,y \in \mathbf{R})$ .  $\xi$  és  $\eta$  *valószínűségi változók függetlenek*, ha együttes eloszlásfüggvényük megegyezik eloszlásfüggvényeik szorzatával. Valószínűségi változók együttes eloszlásfüggvénye mindkét változójában monoton nemcsökkenő, balról folytonos,  $\lim_{x \rightarrow -\infty} F(x,y)=\lim_{y \rightarrow -\infty} F(x,y)=0$ ,  $\lim_{y \rightarrow \infty, x \rightarrow \infty} F(x,y)=1$ ; ha  $a_1 < b_1$  és  $a_2 < b_2$ , akkor  $F(b_1,b_2)-F(a_1,b_2)-F(b_1,a_2)-F(a_1,a_2) \geq 0$ . Bármely, a valószínűségi változók együttes eloszlásfüggvényére vonatkozó ezen feltételeket kielégítő függvényhez létezik két olyan valószínűségi változó, melyek-

nek együttes eloszlásfüggvénye az adott függvény. Az utolsó feltétel  $P(\xi \in [a_1, b_1] \text{ és } \eta \in [a_2, b_2]) \geq 0$ -t jelenti.

### 2.2.1.2.2. Sűrűségfüggvény

Egy adott eseménytéren értelmezett  $\xi: \Omega \rightarrow \mathbf{R}$  függvény *valószínűségi változó*, ha  $\{\omega \mid \xi(\omega) < x\} \in \mathbf{A}$  bármely  $x \in \mathbf{R}$  esetén. A diszkrét valószínűségi változó teljesíti az iménti definíció követelményeit. A definíció a diszkrét valószínűségi változók általánosítása: folytonos esetben  $P(\xi=x)=0$  minden  $x \in \mathbf{R}$  esetén. Valószínűségi változót (ha nem hangsúlyozottan diszkrét) a definíció szerinti értelemben kell érteni.

Egy  $\xi$  valószínűségi változó *folytonos*, ha eloszlásfüggvénye abszolút folytonos. Folytonos valószínűségi változó esetén  $P(\xi=x)=0$  bármely  $x \in \mathbf{R}$  esetén.

Belátható, hogy  $0 \leq P(\xi=x) = \lim_{\Delta x \rightarrow 0} P(x \leq \xi < x + \Delta x) = \lim_{\Delta x \rightarrow 0} (F(x + \Delta x) - F(x)) \leq \lim_{\Delta x \rightarrow 0} ((F(x) + K\Delta x) - F(x)) = \lim_{\Delta x \rightarrow 0} (K\Delta x) = 0$ .

Bármely  $\xi$  valószínűségi változó esetén bármely  $x \in \mathbf{R}$ -re  $P(a \leq \xi \leq b) = F(b+0) - F(a)$ ,  $P(a < \xi < b) = F(b) - F(a-0)$ ,  $P(a < \xi \leq b) = F(b+0) - F(a-0)$  teljesül. Folytonos esetben ezek mind egyenlők  $P(a \leq \xi < b)$ -vel. Folytonos  $\xi$  valószínűségi változó *sűrűségfüggvénye*

$f: \mathbf{R} \rightarrow \mathbf{R}$ , ha  $\int_{-\infty}^x f(t) dt = F(x)$  teljesül minden  $x \in \mathbf{R}$  esetén. Folytonos valószínűségi

változó sűrűségfüggvényére és eloszlásfüggvényére teljesül az  $F'(x) = f(x)$  összefüggés. Diszkrét valószínűségi változónak nincs sűrűségfüggvénye.

A sűrűségfüggvény nemnegatív, az egész számegegyenesen vett integrálja 1. Bármely, a valószínűségi változó sűrűségfüggvényére vonatkozó, ezen feltételeket kielégítő függvényhez létezik olyan valószínűségi változó, melynek sűrűségfüggvénye az adott függvény. Diszkrét esetben hasonló feltétel a  $p_i \geq 0$ ,  $\sum p_i = 1$  tulajdonságok: az eloszlás játssza a sűrűségfüggvény szerepét is.

$\xi$  és  $\eta$  abszolút folytonos *valószínűségi változók együttes sűrűségfüggvénye*

$f(u, v)$ , ha rá teljesül az  $\int_{-\infty}^x \int_{-\infty}^y f(u, v) dv du = F_{\xi, \eta}(x, y)$  egyenlőség.

Az együttes sűrűségfüggvény megegyezik az együttes eloszlásfüggvény második parciális deriváltjával (először az első, majd a második változó szerint deriválva). Abszolút folytonos  $\xi$  és  $\eta$  valószínűségi változók esetén függetlenségük egyenértékű  $f_{\xi, \eta}(u, v) = f_{\xi}(u)f_{\eta}(v)$  teljesülésével. Az együttes sűrűségfüggvény nemnegatív, az egész valós számsíkon vett integrálja 1. Az együttes sűrűségfüggvényből meghatározható az alkotó valószínűségi változók sűrűségfüggvénye. Az együttes sűrűség-

függvényt az egyik változó szerint az egész valós számegyenesen integrálva a másik valószínűségi változó sűrűségfüggvényét kapjuk.

### 2.2.1.3. Állandó és nem állandó véletlen folyamatok

Bár a jel  $x(m)$ -mel jelölt amplitúdója  $m$  periódussal hullámzik, a jel által generált folyamat jellemzője lehet időben állandó vagy időben változó. Nem állandó folyamat például a beszéd, melynek hangossága és spektrális jellemzői, összetevői folyamatosan változnak, ahogyan a beszélő különféle hangokat képez. A folyamat állandónak tekinthető, ha a valószínűségi modellbeli jellemzői idő invariánsak, ellenkező esetben nem állandó – mindkettőt a 12. ábra szemlélteti. Az állandó tulajdonság magába foglalja, hogy a jel összes paramétere, úgy mint középérték, variancia, energiaspektrum összetétele, és a folyamat magasabb rendű momentumai idő invariánsak. A gyakorlatban az állandóságnak különféle fokai ismeretesek. Van olyan, hogy a statisztikák egy halmaza állandó, a többi pedig nem. Például lehet egy véletlen folyamatnak idő invariáns középértéke, de az idő függvényében változó energiája.



12. ábra Kvázistacionárius és nemstacionárius beszédsezmensek

### 2.2.1.3.1. Szigorú értelemben vett állandó folyamatok

Egy  $X(m)$  véletlenszerű folyamat szigorú értelemben véve stacionárius, ha minden megoszlása, és statisztikai paramétere idő invariáns. A szigorú értelemben vett állandóság magában foglalja, hogy az  $n$ -edik rendű eloszlása transláció invariáns minden  $n = 1, 2, \dots$ , vagyis:

$$P[x(m_1) \leq x_1, x(m_2) \leq x_2, \dots, x(m_n) \leq x_n] = P[x(m_1 + \tau) \leq x_1, x(m_2 + \tau) \leq x_2, \dots, x(m_n + \tau) \leq x_n]$$

Az egyenletből a szigorú értelemben vett állandó folyamat statisztikája beleértve a középértéket, a korrelációt és az energiaspektrumot, idő invariáns lesz, ezért:  $E[x(m)] = \mu_x$ ,  $E[x(m)x(m+k)] = r_{xx}(k)$  és  $E[|X(f, m)|^2] = E[|X(f)|^2] = P_{xx}(f)$ , ahol  $\mu_{xx}$ ,  $r_{xx}(m)$  és  $P_{xx}(f)$  rendre az  $x(m)$  jel középértéke, az autokorrelációja és az energiaspektruma, és  $X(f, m)$  jelöli  $x(m)$  frekvencia-idő spektrumát.

### 2.2.1.3.2. Nem szigorú értelemben vett állandó folyamatok

Míg a szigorú értelemben vett állandó (stacionárius) folyamatok esetében a minden statisztikának idő-invariánsnak kell lennie, a nem szigorú értelemben vett stacionárius folyamatokra nem vonatkozik az összes korlátozás. Egy ilyen folyamat középértéke és az autokorrelációs függvénye idő invariáns:  $E[x(m)] = \mu_x$ ,  $E[x(m)x(m+k)] = r_{xx}(k)$ . A két folyamat definíciójából kiderül, hogy minden szigorú értelemben vett stacionárius folyamat nem szigorú értelemben is az, viszont ennek a fordítottja nem igaz.

### 2.2.1.3.3. Nem stacionárius folyamatok

Egy véletlenszerű folyamat nem stacionárius, ha eloszlása vagy statisztikája időben változik. A legtöbb sztochasztikus folyamat, mint például a videojelek, audiojelek, gazdasági adatok, meteorológiai adatok, stb., nem stacionárius, mert egy-egy olyan rendszer generálja őket, melynek környezeti paraméterei időben változnak. Például a beszéd nem stacionárius, mert egy időben változó artikulációs rendszer állítja elő. A beszéd hangerőssége, frekvencia-összetétele időben változó, mely változás néha teljesen váratlan. Az időben változó folyamatok modellezhetők stacionárius véletlenszerű folyamatok egy kombinációjával.

### 2.2.1.4. Véletlen folyamatok várható értéke

A várható érték központi szerepet játszik a jelek modellezésében és feldolgozásában. Továbbá a véletlenszerű folyamatok valószínűségi modelljei általában a várható érték függvényeként kerülnek kifejezésre. Például a normális eloszlás el-



oszlásfüggvénye a folyamat középértékét és kovarianciáját tekintve egy exponenciális függvényként definiálható, a Poisson eloszlás eloszlásfüggvénye pedig a folyamat középértékével definiálható. A jelfeldolgozási alkalmazásokban mindig megvan a jel megfelelő statisztikai modellje – például a normális eloszlás eloszlásfüggvénye –, és a modell kiegészítésére szükség van a várható paraméterek értékére. Azonkívül sok jelfeldolgozási algoritmusban, mint például a spektrális kivonásban a zajcsökkentésre, vagy a lineáris előrejelzésben, lényegében amire szükségünk van az a jel középértékének vagy a korrelációs függvényének kiszámítása. Egy véletlenszerű folyamat  $h(X(m_1), X(m_2), \dots, X(m_M))$  függvényének várható értékének kiszámítása a következőképpen definiálható:

$$E[h(X(m_1), \dots, X(m_M))] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} h(x_1, \dots, x_M) f_{X(m_1) \dots X(m_M)}(x_1, \dots, x_M) dx_1 \dots dx_M$$

A legfontosabb és széles körben használt várható értékek a középérték, a korreláció, a kovariancia és a energiaspektrum.

#### 2.2.1.4.1. A középérték

A egy jelsorozat középértéke nagyon fontos szerepet játszik a jelfeldolgozásban és paraméterkiszámításban a zajos jelek esetében. Például egy zajos jel esetében az optimális lineáris kiszámítási módszer az interpoláció a középérték és a megfigyelt jel zajjal terhelt értéke között. Egy  $[X(m_1), \dots, X(m_M)]$  véletlenszerű vektor középértéke alatt értjük a teljes jelfolyam alatt mutatott átlagértéket, mely a következőképpen definiálható:

$$E[X(m_1), \dots, X(m_M)] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (x_1, \dots, x_M) f_{X(m_1) \dots X(m_M)}(x_1, \dots, x_M) dx_1 \dots dx_M$$

#### 2.2.1.4.2. Autokorreláció

A korrelációs függvény és annak Fourier transzformációja, az energiaspektrum sűrűsége minták és struktúrák modellezésére és felismerésére használatos a jelfeldolgozások során. A kölcsönösségi együtthatók központi szerepet játszanak a jelfeldolgozásban, a telekommunikációs rendszerekben, beleértve a prediktív kódolókat, kiegyenlítőket, digitális dekódereket, késleltetés kiszámítókat, osztályozókat és jelvisszaállító rendszereket. Egy  $X(m)$  véletlenszerű folyamat autokorrelációs függvénye, melyet  $r_{xx}(m_1, m_2)$ -vel jelölünk, a következőképpen definiálható:

$$r_{xx}(m_1, m_2) = E[x(m_1)x(m_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(m_1)x(m_2) f_{X(m_1), X(m_2)}(x(m_1), x(m_2)) dx(m_1) dx(m_2)$$

Az  $r_{xx}(m_1, m_2)$ -vel jelölt autokorrelációs függvény az  $X$  folyamat végeredményének  $m_1$  és  $m_2$  időpillanatbeli hasonlóságának mértékét vagy a kölcsönösségi relációját mutatja meg. Ha a véletlenszerű folyamat végeredménye az  $m_1$  és az  $m_2$  időpillanatbeli értéke nincs hordoz relációt, akkor  $X(m_1)$  és  $X(m_2)$  függetlennek vagy korrelálatlannak mondható, tehát  $r_{xx}(m_1, m_2) = 0$ . A nem szigorú értelemben vett stacionárius folyamatok esetében az autokorrelációs függvény idő invariáns és a két időpillanat között eltelt idő hosszának függvénye, vagyis  $m = m_1 - m_2$ :  $r_{xx}(m_1 + \tau, m_2 + \tau) = r_{xx}(m_1, m_2) = r_{xx}(m_1 - m_2) = r_{xx}(m)$ . Egy valós értékű nem szigorú értelemben vett stacionárius folyamat autokorrelációs függvénye szimmetrikus a következő tulajdonságokkal:  $r_{xx}(-m) = r_{xx}(m)$  és  $r_{xx}(m) \leq r_{xx}(0)$ . Ez esetben az  $r_{xx}(0)$  jelenti a nulla középértékű jel energiáját.

#### 2.2.1.4.3. Autokovariancia

Az autokovariancia függvény egy  $X(m)$  véletlenszerű folyamat  $c_{xx}(m_1, m_2)$  jelöléssel ellátott szórásának vagy szóródásának mérésére szolgál a folyamat középértéke körül és a következőképpen definiálható:

$$c_{xx}(m_1, m_2) = E[(x(m_1) - \mu_x(m_1))(x(m_2) - \mu_x(m_2))] = r_{xx}(m_1, m_2) - \mu_x(m_1)\mu_x(m_2),$$

ahol a  $\mu_x(m)$  jelenti az  $X(m)$  véletlenszerű folyamat középértékét. Nulla középértékű folyamatok esetében az autokovariancia és az autokorreláció függvények azonosak. A  $c_{xx}(m_1, m_2)$  a folyamat varianciáját (szórásnégyzetét) jelenti. Stacionárius folyamatok esetében az autokovariancia függvény az előző egyenletből a következőképpen alakul:  $c_{xx}(m_1, m_2) = c_{xx}(m_1 - m_2) = r_{xx}(m_1 - m_2) - \mu_x^2$ .

#### 2.2.1.4.4. Spektrális energiasűrűség

Egy véletlenszerű folyamat energiájának spektrális sűrűségfüggvénye, más néven energiaspektruma megadja a jel egyes frekvenciákhoz tartozó energiáinak eloszlását a teljes spektrumon. A nem szigorú értelemben vett stacionárius folyamat  $X(m)$  energiaspektruma a Wiener–Khinchin tétel alapján úgy definiálható, hogy az autokorrelációs függvény Fourier transzformáltja és a következőképpen írható fel:  $P_{XX}(f) = E[X(f)X^*(f)] = \sum_{m=-\infty}^{\infty} r_{xx}(m)e^{-j2\pi fm}$ , ahol  $r_{xx}(m)$  és  $P_{XX}(f)$  rendre az  $x(m)$  autokorrelációs függvénye és energiaspektruma, és  $f$  pedig a frekvenciaváltozó. Egy valós értékű stacionárius folyamat autokorrelációs függvénye szimmetrikus, és az energiaspektruma a következőképpen írható fel:

$$P_{XX}(f) = r_{xx}(0) + \sum_{m=1}^{\infty} 2r_{xx}(m) \cos(2\pi fm)$$

A spektrális sűrűségfüggvény egy nemnegatív valós értékészletű felvevő függvény, megadása pedig watt per herzben  $\left[\frac{\text{W}}{\text{Hz}}\right]$  történik. Az autokorrelációs so-

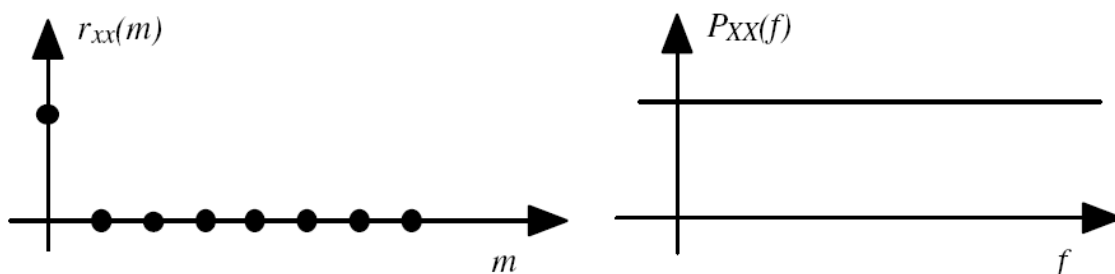
rozat a véletlenszerű folyamat energiaspektrumából az iménti egyenlet felhasználásával kapható meg az inverz Fourier transzformáció alkalmazásával a követke-

zőképpen:  $r_{xx}(m) = \int_{-\frac{1}{2}}^{\frac{1}{2}} P_{XX}(f) e^{j2\pi fm} df$ . Az autokorreláció és az energiaspektrum a

folyamat másodrendű statisztikai, rendre időben és frekvenciatartományban.

#### Példa – fehér zaj

A következő ábra a fehér zaj energiaspektrumát és autokorrelációját mutatja.



13. ábra A fehér zaj autokorrelációja és energiaspektruma

Egy olyan zajt, amely korrelálatlan, egymástól független mintákból áll, fehér zajnak nevezünk. Egy stacionárius fehér zaj  $n(m)$  autokorrelációja definiálható:

$r_{nn}(k) = E[n(m)n(m+k)] = \begin{cases} \text{Zajenergia} & k = 0 \\ 0 & k \neq 0 \end{cases}$ . Ez az egyenlet egy korrelálatlan fehér

zaj definíciója szerinti matematikai állítás. A frekvenciatartományra ekvivalens leírás adható meg, az  $r_{nn}(k)$  Fourier transzformációjával a következőképpen:

$P_{NN}(f) = \sum_{k=-\infty}^{\infty} r_{nn}(k) e^{-j2\pi fk} = r_{nn}(0) = \text{zajenergia}$ . A stacionárius fehér zaj energiaspekt-

ruma azonosan terjed szét a sáv teljes szélességében és időben. A fehér zaj az egyik legnehezebben eltávolítható zajforma, mert nem határozható meg a szerkezete sem időben, sem a frekvenciasávok szerint.

#### Példa – impulzív zaj

Az impulzív zaj egy véletlenszerű kétállapotú („be/ki”) impulzussorozatból áll, véletlenszerű amplitúdóval és időben véletlenszerű előfordulással. Egy véletlenszerű impulzív zaj sorozat  $n_i(m)$  egy kétállapotú véletlenszerű amplitúdó modulált sorozatként modellezhető, a következő szerint:  $n_i(m) = n(m)b(m)$ , ahol  $b(m)$  egy

kétállapotú sorozat, mely egy impulzus jelenlétét vagy hiányát jelzi, és  $n(m)$  pedig egy véletlenszerű zaj. Tegyük fel, hogy az impulzív zaj egy korrelálatlan folyamat, így az impulzív zaj, mint egy kétállapotú folyamat autokorrelációja a következőképpen definiálható:  $r_m(k, m) = E[n_i(m)n_i(m+k)] = \sigma_n^2 \delta(k)b(m)$ , ahol  $\sigma_n^2$  a zaj varianciája. Ebben az egyenletben az autokorrelációs függvény egy kétállapotú függvényként lett kifejezve, mely az impulzív zaj  $m$  időpillanatbeli be/ki állapotától függ. Az impulzív zaj sorozat energiaspektruma az imént leírt autokorrelációs függvény Fourier transzformáltjaként kapható meg a következő összefüggéssel:  $P_{NN}(f, m) = \sigma_n^2 b(m)$ .

#### 2.2.1.4.5. Két véletlenszerű folyamat együttes statisztikája

A jelfeldolgozás során gyakran találkozunk azzal a problémával, például egy több érzékelőből álló rendszer esetében, hogy nem csak egy véletlenszerű folyamat van jelen. Az együttes statisztika és együttes eloszlás leírja a két vagy több folyamat kapcsolatát egymással. Két diszkrét időbeosztással mintavételezett véletlenszerű folyamat  $x(m)$  és  $y(n)$  együttes eloszlásfüggvénye a következő jelöléssel adható meg:  $f_{X(m_1)\dots X(m_M), Y(n_1)\dots Y(n_N)}(x_1, \dots, x_M, y_1, \dots, y_N)$ . Ha a két véletlenszerű folyamat  $X(m)$  és  $Y(n)$  korrelálatlanok, az együttes eloszlásfüggvényük a két eloszlásfüggvény szorzataként adható meg a következő felírással:

$$f_{X(m_1)\dots X(m_M), Y(n_1)\dots Y(n_N)}(x_1, \dots, x_M, y_1, \dots, y_N) = f_{X(m_1)\dots X(m_M)}(x_1, \dots, x_M) f_{Y(n_1)\dots Y(n_N)}(y_1, \dots, y_N)$$

#### 2.2.1.4.6. Kereszt-korreláció és kereszt-kovariancia

Két véletlenszerű folyamat kereszt-korrelációja az alábbiak szerint adható

$$\text{meg: } r_{xy}(m_1, m_2) = E[(x(m_1)y(m_2))] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(m_1)y(m_2)f_{X(m_1)Y(m_2)}(x(m_1), y(m_2))dx(m_1)dy(m_2)$$

Egy  $r_{xy}(m_1, m_2)$  nem szigorú értelemben vett stacionárius folyamat esetében a korrelációs függvény csak az  $m_1$  és  $m_2$  időpillanatok különbségétől függ, vagyis  $m = m_1 - m_2$ , tehát:  $r_{xy}(m_1 + \tau, m_2 + \tau) = r_{xy}(m_1, m_2) = r_{xy}(m_1 - m_2) = r_{xy}(m)$ .

A kereszt-kovariancia függvény a következőképpen definiálható:  $c_{xy}(m_1, m_2) = E[(x(m_1) - \mu_x(m_1))(y(m_2) - \mu_y(m_2))] = r_{xy}(m_1, m_2) - \mu_x(m_1)\mu_y(m_2)$ . Nulla középértékkel rendelkező folyamatok esetében a kereszt-korreláció és a kereszt-kovariancia függvények azonosak. Nem szigorú értelemben vett stacionárius folyamatok esetében a kereszt-kovariancia függvény az előző egyenlet alapján:  $c_{xy}(m_1, m_2) = c_{xy}(m_1 - m_2) = r_{xy}(m_1 - m_2) - \mu_x\mu_y$ .

## 2.2.1.4.7. Kereszt-energiaspektrum eloszlás és koherencia

Két véletlenszerű folyamat  $X(m)$  és  $Y(m)$  kereszt-energiaspektrum eloszlása a két folyamat kereszt-korrelációs függvényének Fourier transzformációjával definiálható:

$$P_{XY}(f) = E[X(f)Y^*(f)] = \sum_{m=-\infty}^{\infty} r_{xy}(m)e^{-j2\pi fm}.$$

Ahogy a kereszt-korreláció a kereszt-energiaspektrum eloszlása is a két jel hasonlóságának mértékét mutatja, vagy a két folyamat spektrális koherenciáját, mely a kereszt-energiaspektrum eloszlásának normalizált formulája és a következőképpen definiált:

$$C_{XY}(f) = \frac{P_{XY}(f)}{\sqrt{P_{XX}(f)P_{YY}(f)}}.$$

A koherencia függvény olyan alkalmazásokban használatos, mint az időkésettetés kiszámítása és a jel-zaj arány meghatározása (mérése).

## 2.2.2. A rejtett Markov modell

A rejtett Markov modelleket a véletlenszerű nem stacionárius jelek statisztikai modellezésére használhatjuk hatékonyan. Az ilyen jelek csoportjába tartozik az emberi beszéd vagy az időben változó zaj. A HMM egy Markov láncot ad meg, mely az időbeli változást vagy a terek változását állapotfüggő stacionárius alfolyamatokra bontja. Egy HMM lényegében egy véges állapothalmazzal rendelkező Bayes folyamat, Markov előzménnyel, az egyes állapotok között átmenetek modellezésére, és egy halmaz valószínűségi eloszlásfüggvény a jel véletlenszerű variációinak modellezésére minden állapoton belül.

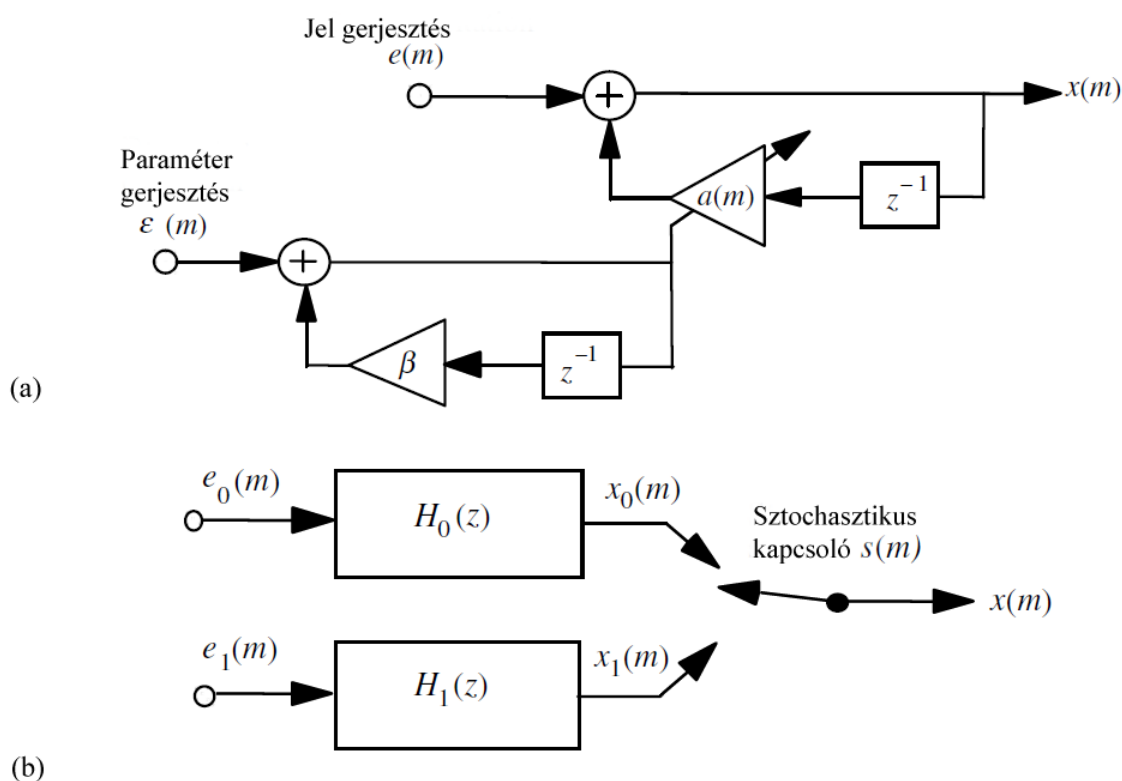
2.2.2.1. Nem stacionárius folyamatok statisztikai modellje

Egy nem stacionárius folyamat definiálható oly módon, hogy egy olyan folyamat, melynek statisztikai paraméterei az idővel változnak. A legtöbb „természetesen előállított” jel, mint például a hangjelek, képek, biomedikai és szeizmikus jelek nem stacionárius jelek, mivel az azokat előállító rendszerek paraméterei és a környezet, melyben terjednek változnak az idő múlásával.

A nem stacionárius jelek egy kétrétegű sztochasztikus folyamatként modellezhetők, egy rejtett folyamattal az észlelhető jel statisztikai paramétereinek időbeli változásainak kontrollálására. Általában a nem stacionárius folyamatok két alapvető nagy csoportba sorolhatók:

- a.) folytonosan változó állapotú folyamatok
- b.) véges sok állapotú folyamatok

A folytonosan változó állapotú folyamatok definíciós szerint az időben folytonosan változó jelek statisztikájának alapjául szolgálnak. Ilyenek például az audiojelek, a beszéd és a zene, melyeknek az energiája és spektrális összetétele az idővel változik. Véges sok állapottal rendelkező folyamat az olyan statisztikai jellemzőkkel rendelkező folyamat, mely képes váltani a véges sok stacionárius vagy nem stacionárius állapot között. Például az impulzív zaj egy ilyen kétállapotú folyamat. Folytonosan változó folyamatok közelíthetők hozzájuk illő véges sok állapotú folyamattal. A következő ábra egy nem stacionárius autoregresszív (AR) folyamatot mutat be. Ez a folyamat a jel paramétereire egy rejtett stacionárius AR modell és magára a jelre egy időben változó AR modell kombinációját alkalmazza. A



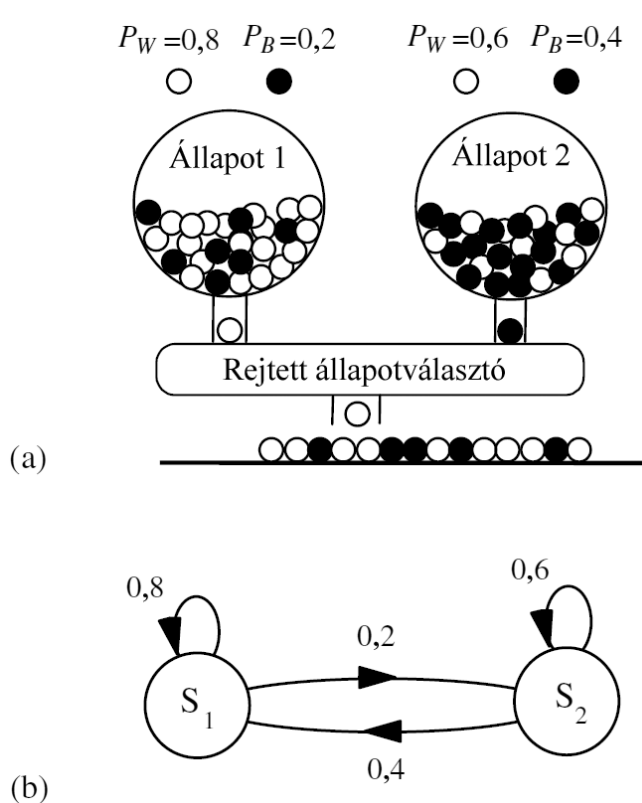
14. ábra Folytonosan változó (a) és kétállapotú (b) AR folyamat

rejtett modell vezérli a nem stacionárius AR modell időben változó paramétereit. Ehhez a modellhez a megfigyelt jel egyenlete és a paraméter állapotegyenlete a következő módon írható fel: a megfigyelt jel  $x(m) = a(m)x(m-1) + e(m)$  és a rejtett állapot  $a(m) = \beta a(m-1) + \varepsilon(m)$ , ahol  $a(m)$  a megfigyelt AR folyamat időben változó együtthatója és  $\beta$  pedig a rejtett állapotirányító folyamat együtthatója. A 14. (b) ábra egy egyszerű példát mutat a véges sok állapotú nem stacionárius folyamat kétállapotú autoregresszív megvalósítására, amelyben egy közvetlen kapcsoló választja ki minden időpillanatban a két AR modell közül az egyiket, melyet a

kimenetre csatol. Ehhez a modellhez a kimeneti jel  $x(m)$  a következő egyenlettel fejthető ki:  $x(m) = \bar{s}(m)x_0(m) + s(m)x_1(m)$ , ahol a kétállású kapcsoló  $s(m)$  kiválasztja a folyamat állapotát  $m$  időpillanatban, és  $\bar{s}(m)$  jelöli  $s(m)$  ellentettjét.

### 2.2.2.2. Rejtett Markov modellek

A rejtett Markov modell (HMM) egy kétrétegű véges sok állapotú folyamat, egy rejtett Markov folyamattal, mely irányítja az állapotok kiválasztását a figyelembe veendő folyamatban. Mint egy egyszerű illusztrációt figyeljük meg a következő ábrán látható kétállapotú Markov folyamatot, mely két tárolót mutat különböző keverékű fehér és fekete színű golyókkal. A fekete és fehér golyók előfordulásának valószínűségét jelöljük rendre  $P_B$  és  $P_W$  jelölésekkel, ahogyan az ábra mutatja. Tegyük fel, hogy folytonos idő intervallumban egy rejtett kiválasztó folyamat kivá-



15. ábra Kétrétegű véletlenszerű folyamat (a), és rejtett Markov modellje (b)

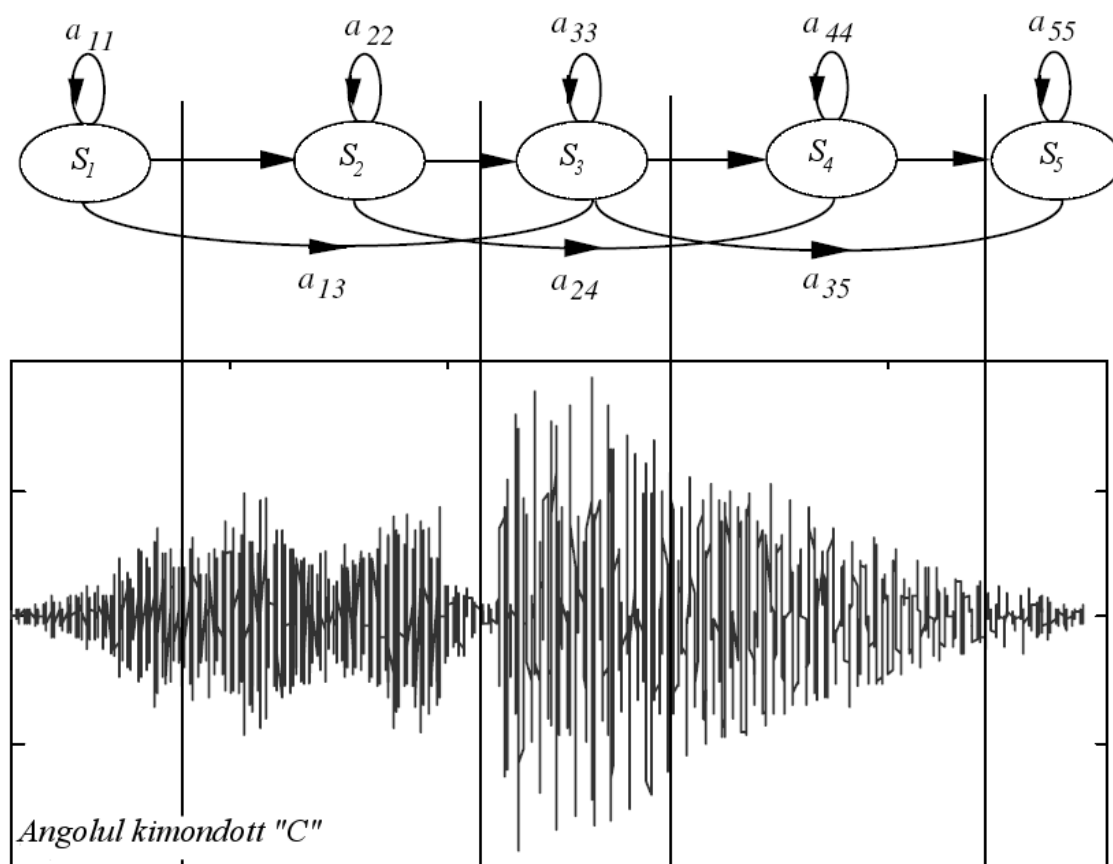
lasztja az egyik tartályt, hogy kiejtsen egy golyót. A golyót kiejtése után pótoljuk, tehát a fehér és fekete golyók számára nincs hatása a műveletnek. Minden tárolót úgy tekinthetünk, hogy az egy alapvető állapota a kimeneti folyamatnak. Most tegyük fel, hogy a rejtett tároló-kiválasztót a következő szabály irányítja: bármely időpillanatban, ha a kiválasztott tárolóból fehér golyót kapunk, akkor nem választunk másik tárolót, ha feketét akkor a másik tárolót választjuk. Ez egy példa a Markov folyamatra, mert a folyamat következő állapota függ a jelenlegi állapottól, ahogyan azt az ábra kétállapotú modellje is mutatja. Ebben a modellben a megfigyelhető végeredmény nem jelzi egyértelműen az alapul vett állapotot, mivel mindkét állapot (tároló) alkalmas minkét színű (fekete és fehér) golyó kibocsátására.

Általában a Markov modellnek  $N$  állapota van, és mindegyik állapot más szegmensét modellezi a vizsgált jelnek. A rejtett Markov modell alkalmazható egy

időben változó véletlenszerű folyamat modellezésére, mint egy  $N$  elemű stacionárius vagy kvázi-stacionárius elemi alfolyamat valószínűségi Markov lánc.

### 2.2.2.2.1. A rejtett Markov modellek fizikai értelmezése

A modell fizikai értelmezéséhez nézzük meg a következő, 16. ábrát. Az ábra balról jobbra mutatja egy angolul kiejtett „C”<sup>1</sup> hang modelljét és vele együtt a hang hullámformájának térképét is. Általában kétféle eltérés van a beszéd és az egyéb sztochasztikus jelek között: a spektrális összeállítás változása, és az időbeosztás vagy artikuláció változása. Egy rejtett Markov modellben ezek a változások állapot megjegyzéssel és az egyes állapotok közötti átmenetek valószínűségével modellezhetők. Használható módszer a rejtett Markov modell interpretálására és alkalmazására, hogy a modell minden állapotát úgy tekintjük, hogy a sztochasztikus folyamat egy szegmensét modellezi. Az ábrán az  $S_1$  állapot modellezi az első szeg-



16. ábra Ötállapotú balról–jobbra HMM beszédmodell

<sup>1</sup> Az angol abc 3. betűjének, a c-nek kiejtése a magyartól jelentősen eltér, fonetikusan a következőképpen írható le: [si:]



menet, a  $S_2$  a másodikat és így tovább. Minden állapotnak kell lennie egy olyan mechanizmusának, mely összeilleszti a szegmensek véletlenszerű variációinak különböző realizációit, melyet modelleznek. Az állapotátmenetek valószínűségei egy mechanizmust biztosítanak a különböző állapotok kapcsolatára, és a variációk modellezésére az időskála jelek alatti időtartamára minden állapotban. Például ha a beszéd egy szegmensének vége kinyúlt, köszönhetően a lassú artikulációnak, hozzáigazítható lesz több önmagába visszatérő átmenettel az állapothoz, mely a szegmenst modellezi. Ellenben ha a szegmens elhallgatott a gyors beszédnek köszönhetően, akkor a következő állaputra ugrással illeszt ebben az esetben. Az állapot tapasztalati eloszlásfüggvényei modellezi a jel állapotokhoz rendelt szegmenseinek spektrális összeállításának valószínűségi eloszlását.

#### 2.2.2.2.2. A rejtett Markov modell paraméterei

A rejtett Markov modellnek a következő paraméterei vannak:

- Az állapotok száma  $N$ . Ezt általában beállítjuk az egymástól különböző vagy elemi sztochasztikus események számára egy jelfolyamatban. Például egy impulzív zajhoz hasonló kétállapotú folyamat modellezése esetén  $N$  értéke 2 lesz, míg szavankénti beszéd esetében  $N$  értékét 5 és 10 közé szokás választani.
- Állapot-átmeneti valószínűségi mátrix  $\mathbf{A} = \{a_{ij}, i, j = 1, \dots, N\}$ . Ez biztosít kapcsolatot vagy Markov hálózatot hoz létre az egyes állapotok között, és modellezi a variációkat az állapotokhoz rendelt jelek időtartama alatt. A balról–jobbra rejtett Markov modell esetében  $a_{ij} = 0$ , minden  $i > j$  esetén, ezért ez az  $\mathbf{A}$  mátrix egy felső háromszög mátrix.
- Állapot tapasztalati vektorok  $\{\mu_{i1}, \mu_{i2}, \dots, \mu_{iM}, i = 1, \dots, N\}$ . Minden állapotra egy halmaz  $M$  prototípus vektor, a jel terének centroidjainak modellezésére minden állapotban.
- Állapot tapasztalati vektor valószínűségi modellje. Lehet diszkrét az  $M$  prototípusvektorokból, azok hozzárendelt sűrűségfüggvényével  $\mathbf{P} = \{P_{ij}(\cdot); i=1, \dots, N, j=1, \dots, M\}$ , vagy folytonos (általában normális)  $\mathbf{F} = \{f_{ij}(\cdot); i=1, \dots, N, j=1, \dots, M\}$  eloszlásfüggvény modell.
- Kezdőállapot valószínűségi vektor  $\boldsymbol{\pi} = [\pi_1, \pi_2, \dots, \pi_N]$

2.2.2.3. A rejtett Markov modell betanítása

Az első lépés, hogy adatbázist alakítsunk ki a modell számára, nagy számú minta vételezésével a vizsgálat tárgyát képező jelből. Tegyük fel, hogy az adatbázis  $L$  vektor értékekkel bíró szekvenciából áll  $[\mathbf{X}]=[\mathbf{X}_k; k=0, \dots, L-1]$ , és minden szekvenciában  $\mathbf{X}_k=[x(t); t=0, \dots, T_k-1]$  változó számú  $T_k$  vektor van. A paraméterek megfelelő beállításához a modell statisztikájának adathalmazát kell előállítani. Valószínűségi értelemben a modell alkalmassága a modell utólagos valószínűségével  $P_{M|X}(M|\mathbf{X})$  adható meg,  $M$  modell  $\mathbf{X}$  adathalmazára. Ez a betanító folyamat az  $M$  modell utólagos valószínűségének és az  $[\mathbf{X}]$  adathalmaznak a maximalizálását célozza, a Bayes szabály alkalmazásával:  $P_{M|X}(M|\mathbf{X}) = \frac{1}{f_X(\mathbf{X})} f_{X|M}(\mathbf{X}|M)P_M(M)$ ,

ahol a nevezőben álló  $f_X(\mathbf{X})$  egy normalizáló hatással bír, és  $P_M(M)$  az  $M$  modell korábbi valószínűsége. Az iménti egyenlet maximalizálása egyenértékű a valószínűségi függvény maximalizálásával, mely a következő egyenlet segítségével fejezhető ki  $\mathbf{X}$  vektorszekvenciára és  $M$  modellre:

$$f_{X|M}(\mathbf{X}|M) = \sum_s f_{X|S,M}(\mathbf{X}|s, M)P_{s|M}(s|M),$$

ahol  $f_{X|S,M}(\mathbf{X}|s, M)$  jelöli az eloszlásfüggvényt az  $M$  modell  $\mathbf{X}$  szekvenciájának  $\mathbf{s}=[s(0), s(1), \dots, s(T-1)]$  állapotában, mely a következőképpen fejezhető ki:

$$f_{X|S,M}(\mathbf{X}|s, M) = f_{X|S}(\mathbf{x}(0)|s(0))f_{X|S}(\mathbf{x}(1)|s(1))\cdots f_{X|S}(\mathbf{x}(T-1)|s(T-1)),$$

ahol az  $N$  állapot egyike lehet  $s(t)$   $t$  időpillanatban,  $f_{X|S}(\mathbf{x}(t)|s(t))$  pedig

$f_{X|S,M}(\mathbf{x}(t)|s(t), M)$  rövidített leírása az  $M$  modell  $\mathbf{x}(t)$  eloszlásfüggvényének  $s(t)$

állapotban. A  $\mathbf{s}$  állapotszekvencia Markov valószínűsége a következőképpen adható

$$\text{meg: } P_{s|M}(s|M) = \pi_{s(0)}a_{s(0)s(1)}a_{s(1)s(2)}\cdots a_{s(T-2)s(T-1)}.$$

Az előző három egyenletből az utolsó kettőt behelyettesítve az elsőbe kapjuk a következőt:

$$\begin{aligned} f_{X|M}(\mathbf{X}|M) &= \sum_s f_{X|S,M}(\mathbf{X}|s, M)P_{s|M}(s|M) = \\ &= \sum_s \pi_{s(0)}f_{X|S}(\mathbf{x}(0)|s(0))a_{s(0)s(1)}f_{X|S}(\mathbf{x}(1)|s(1))\cdots a_{s(T-2)s(T-1)}f_{X|S}(\mathbf{x}(T-1)|s(T-1)) \end{aligned}$$

ahol az összegzés végighalad a összes  $\mathbf{s}$  állapotszekvenciát. A betanítási folyamatban az átmeneti valószínűségeket és a tapasztalati eloszlásfüggvények paramétereit az úgy kell kiszámítani, hogy iménti egyenletből a modell valószínűsége maximális legyen. A közvetlen maximalizálás a modell paramétereivel összefüggésben nem triviális megoldás. Következésképpen az egyenlet számításigénye egy megfigyelési

---

szekvencia  $T$  darab vektorára  $O(N^T)$  lesz. Ez még kis számok esetében is nagyon nagy számításigényt jelent. Az állapot–idő diagram ismétlődő struktúrái azt mutatják, hogy sok az ismétlődő számítás az egyenletben, melyet egy hatékony megvalósítással el tudunk kerülni.

### 3. Zajsűrési módszerek

#### 3.1. Zajzár

Érdekes, bár nem túl hatékony megoldás a jelszint folyamatos vizsgálata, és ha az elér egy kellően alacsony értéket, akkor a jelet teljesen kikapcsolva, csendet kapunk eredményül. Ezzel nem csak a zajt állítjuk meg, hanem az alacsony jelszintű hasznos jelet is, és a jel zajterhelése továbbra is megmarad, ha már egy adott szint felett van az értéke. Ez a módszer megoldást jelenthet olyan esetekben, mikor egy berendezés öngerjesztő folyamatának szeretnénk gátat szabni. Ilyen előfordulhat például mikrofonok esetén, melyek egy légtérben vannak az erősített jelüket megszólaltató hangsugárzókkal, és elhelyezéssel már nem tudunk a problémán segíteni (nagy hangerőre van szükség, és bárhová helyezzük a mikrofont, egy zárt térben, beindul az oszcilláció). Másik szemléletes példa az elektromos gitárok torzító berendezése, melynek alapvető működéséhez tartozik az átlagosnál nagyobb erősítés. Ilyenkor az érkező jel kikapcsolása a húrok öngerjesztő rezgésének megakadályozását szolgálhatja. Egy ilyen berendezés megfelelő beállítása nehéz, főleg, ha több vagy többféle hangforrás is van, ráadásul a valódi problémára nem jelent megoldást.

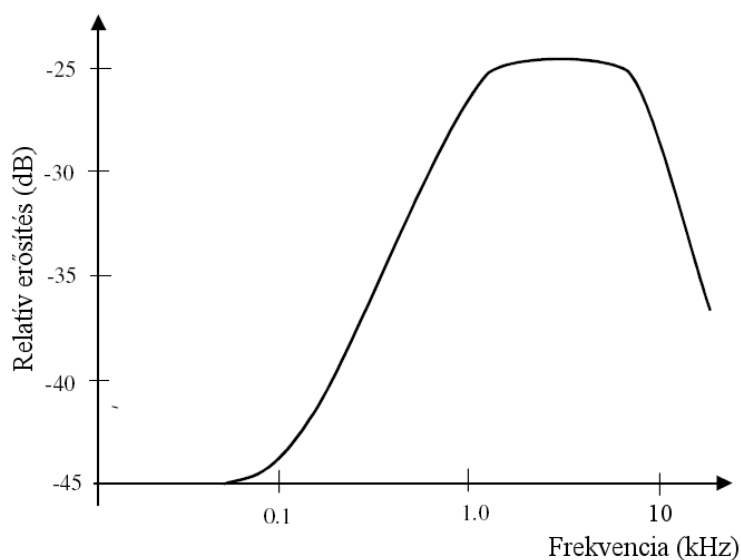
#### 3.2. Analóg valós idejű zajsűrők

##### 3.2.1. Dolby zajsűrő

A Dolby zajsűrő rendszerek úgy működnek, hogy emelik a magasabb frekvenciákon a hasznos jel amplitúdóját, ezzel javítva a jel–zaj arányt ezen frekvenciákon. A jel nagy része főként az alacsonyabb frekvenciákra – 2 kHz alá – koncentrálódik. A magasabb frekvenciájú jelek, melyek a minőséget és az élményt hordozzák viszonylag alacsonyabb energiájúak, és tovább degradálódnak a zajjal. Például egy mágneses szalagra rögzített jelsorozat esetében a szalag zaja és a korrekciós előerősítő fokozat zaja jelentős magasabb frekvenciájú „sistergést” állít elő. Lejátszásakor ezért ezeken a frekvenciákon romlik a jel–zaj arány – tehát nem lesz egyenletes a teljes átvitt sávban, vagyis az alacsonyabb frekvenciákon jobb, a magasabbakon rosszabb értéket kapunk. Tehát a magasabb frekvenciákon kevesebb hallható hasznos jelet kapunk. A Dolby zajsűrő tehát azon az alapelven működik, hogy a viszonylag kisebb energiájú magasabb frekvenciájú jeleket kiemeli, erősíti a hanganyag felvétele során, a többi jelet pedig változatlanul hagyja. Felvétel közben a jel

egy kiemelés és egy dinamika kompresszió keverékén megy keresztül, melyet lejátszáskor vissza kell alakítani az eredeti állapotába. A kódoló és a dekódoló rendszernek jól illesztettnek kell lennie, a torzítások elkerülése érdekében.

A Dolby laboratóriumban többféle ilyen szűrőt fejlesztettek ki, melyek a Dolby A, Dolby B, és Dolby C. Ezek alapjában véve a sávok számában és az alkalmazott kiemelő stratégiájában különböznek egymástól. A Dolby A a professzionális felhasználás számára lett kifejlesztve, és négy nagy részre osztja a frekvenciaspektrumot – az első sáv 0–80 Hz (alul áteresztő szűrő), a második sáv 80–3000 Hz (hullámsávszűrő), a harmadik sáv 3 kHz körüli felül áteresztő szűrő, a negyedik sáv pedig egy 9 kHz körüli felül áteresztő szűrő segítségével jött létre. A kódoló minden hullámsáv erősítését alkalmazkodóan igazította, az alacsony energiájú jelek kiemelése érdekében. A Dolby A rendszer gondoskodik a 10–15 dB-es erősítésről minden sávban, ha a jel 45 dB-nél nagyobb értékkel esik a maximális rögzítési jelszint alá. A kereskedelmi forgalomba kerülő készülékekben a Dolby B és a Dolby C zajsűrőket alkalmazták. Ezek a rendszerek az imént ismertetett Dolby A rendszerrel szemben már csak két hullámsávon dolgoznak, a négy helyett. A



17. ábra A Dolby C rendszer elő-kiemelő sávátvitele

Dolby B rendszer 10 dB-es emelést biztosít alacsony jelszint esetén (nagyobb, mint 45 dB eltérés a maximális rögzítési jelszinthez képest), a Dolby C rendszer pedig 20 dB-es emeléssel dolgozik, ahogy az a következő ábrán látható.

Ez a módszer csak analóg jelek esetében egy jól meghatározott zajforma csökkentésére alkalmas,

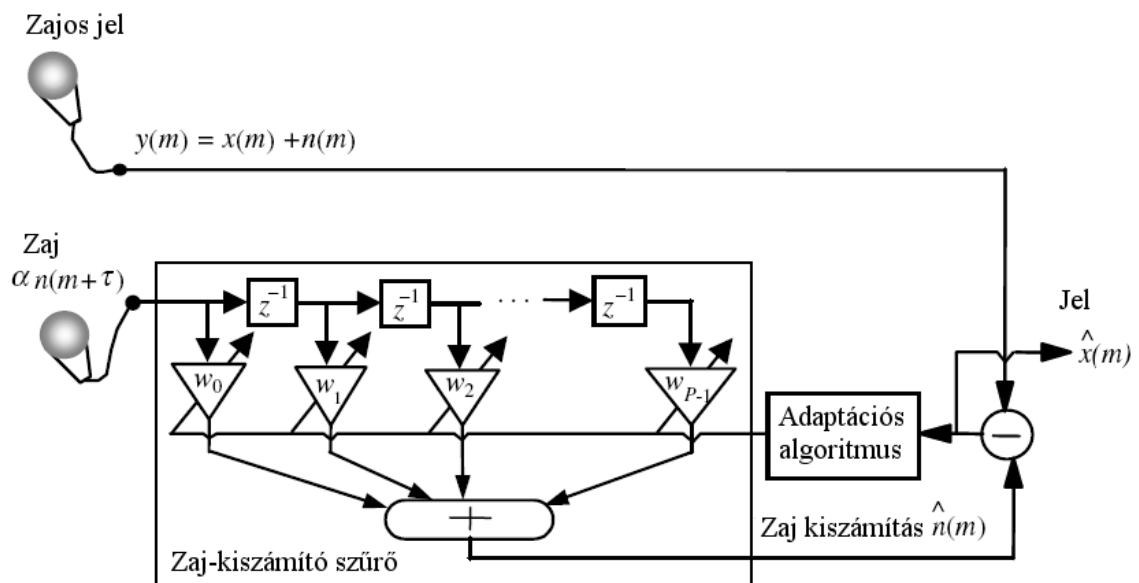
ezért ez nagyon behatárolja az alkalmazási területet. Ezért alkalmazzák ezt a módszert szalagos rögzítő eszközök – úgy mint kompakt kazettás magnetofonok – esetében. Az ilyen készülékek esetében jellemző, hogy a szalagon található mágneses részecskék rendezettsége és az előmágnesezés milyensége befolyásolja a hangminőséget, továbbá a magasabb frekvenciájú hangjelek – az alkalmazott mágne-

szehető anyag függvényében – kisebb jelszinten érkeznek vissza a szalagról, mely egy korrekciós előerősítő fokozat alkalmazását teszi szükségessé, ezáltal az elektronikus rendszernek is lesz egy zaja, melyek együttesének „szűrésére” alkalmazható a Dolby A, B és C rendszer.

### 3.2.2. Zajérvénytelenítés

Egy zajos környezetből (mint például egy mozgó autó) származó beszéd esetén a hasznos jelet a környezet zajával együtt észleljük. A mérőrendszerek esetében az információhordozó jel gyakran a berendezést körülvevő környezet zajával szennyezett. A zajjal együtt észlelt  $y(m)$  jel a következőképpen modellezhető:  $y(m) = x(m) + n(m)$ , ahol  $x(m)$  jelenti a hasznos jelet,  $n(m)$  jelenti a zajt,  $m$  pedig a diszkrét idő mérőszámot. Néhány esetben, például mikor mobiltelefont használunk egy mozgó autóban, vagy rádiókommunikációs eszközt egy repülőgép pilótafülkéjében, talán lehetséges, hogy kimérhető és kiszámítható a környezeti zaj pillanatnyi amplitúdója, egy irányított mikrofon használatával. Ezután a az  $x(m)$  hasznos jel visszaállítható lehet a zajos jel és a zaj különbségeként.

Az adaptív zaj-érvénytelenítő rendszer – két bemenettel rendelkezik, melyekből az egyik a környezeti zajt érzékeli, a másik pedig a zajjal terhelt hasznos jelet – alkalmazható beszédhang kiemelésre. Ebben a rendszerben az egyik irányított mikrofon a zajos  $x(m) + n(m)$  jelet érzékeli, míg a másik irányított mikrofon az előzőtől egy adott távolságra elhelyezve a zajt  $\alpha \cdot n(m + \tau)$  méri. Az a csillapítási szor-



18. ábra Adaptív zajérvénytelenítő elvi felépítése

zótényező és a  $\tau$  időkéleltetési érték meglehetősen túlegyszerűsített modellt ad a mikrofon elhelyezésétől eltérő helyről származó és eltérő zajok hatásainak modellezésére. A második mikrofonból érkező zajt egy adaptív digitális szűrővel dolgozzuk fel, annak érdekében, hogy amplitúdója megegyezzen az első mikrofonban érzékelt hang zajszennyezésének amplitúdójával, ezután pedig ezt a jelet kivonjuk zajos jelből, a zaj érvénytelenítése céljából. Ez a berendezés a zaj alacsonyabb frekvenciáin működik hatékonyan, de szenved a jelek nem állandó jellegétől, és a túlegyszerűsített feltevéstől mely szerint lineáris szűrővel modellezhető a zaj szóródása és terjedése a térben. A berendezés felépítését a 18. ábra szemlélteti.

Sok esetben, például egy telekommunikációs rendszerben alkalmazott vevőkészülék esetében nem ismerjük a zajszennyezés pillanatnyi értékét, csak a zajos jel adott. Ezen esetekben nem érvényteleníthető a zaj, csak mértéke csökkenthető, a jel és a zaj statisztikáját felhasználva. A Wiener szűrő alkalmas ilyen jellegű zajcsökkentési eljárásra, ahol csak a zajos jel adott.

### **3.3. Digitális jelek feldolgoása**

#### *3.3.1. Átalakítások*

##### 3.3.1.1. Analóg–digitális átalakítás

Az analóg jelfeldolgozás során valamilyen folytonosan változó fizikai mennyiséget alakítunk elektronikusan feldolgozható jellé. Ilyen átalakítás történik például egy mikrofonban, ahol a levegő rezgéseit, nyomásváltozásait alakítjuk elektromos feszültséggé, melyet a későbbiekben elektronikus jelerősítő berendezésekkel tudunk tovább feldolgozni. Ezek a jelek még elég sérülékenyek, könnyen terhelődhetnek zajjal, és ezen zajok eltávolítása az analóg jelekből nagyon nehéz feladat. Ha a jelet szeretnénk megszabadítani az additív zajoktól, akkor az első lépés, amit meg kell tennünk a jel átalakítása digitális jellé. A digitális jel fontos jellemzője, hogy egy előre meghatározott értékhalmból vehet fel csak értékeket. Ezen jelek esetében a feldolgozás már matematikai modellek alapján digitális számítógépek segítségével történhet. Továbbá nagyon előnyös tulajdonságaik közé tartozik, hogy sokkal kevésbé érzékenyek a külső additív zajokra, a diszkrét értékek miatt. A legegyszerűbb ilyen digitális jelek kétértékűek, vagy más néven bináris jelek. A mai korszerű kommunikációs rendszerekben gyakran alkalmaznak olyan készülékeket, melyek az analóg jelet először csak digitálissá alakítják, majd ez a digitális jelet

alakítják tovább egy bitfolyammá, melynek értelmezése a vevő oldalon történik. Ez a bit soros adatátvitel, napjaink egyik legkedveltebb átviteli formája.

Az analóg–digitális átalakítás vagy konverzió (Analog to Digital Conversion – ADC) úgy történik, hogy egy analóg feszültségjelet alapul véve abból adott időközönként mintát veszünk. Ez a digitális jel egyik fontos jellemzője, a *mintavételezési frekvencia*. A legjobb minőségű digitális jelet akkor kapnánk, ha egy időegység alatt végtelen sokszor vennénk mintát az analóg jelből, de ez technikailag kivitelezhetetlen és értelmetlen is. A mintavételezés a gyakorlatban rövid időközönként történik:

$T_s = \frac{1}{F_s}$ , ahol  $T_s$  az időközt,  $F_s$  pedig a frekvenciát jelöli. Ez a *mintavételezési eljárás*

az  $x(t)$  analóg jelet egy  $x(m)$  impulzussorozattá alakítja.

Az átalakítás következő lépése a kvantálás, amikor a mintavételezés után kapott impulzusokat egy diszkrét értékhalmozra képezzük. A halmaz elemeit egy analóg jeltartomány meghatározott számú részekre osztásával kapjuk meg. A kvantálás során a mintavételezés eredményeképpen kapott impulzussorozat eleméről eldöntjük, hogy melyik két érték közé (melyik tartományba) esik, és előállítjuk a  $x(m)$  jelből az  $y(m)$  jelet, mely a kvantálás eredményeképpen keletkezik. Ez a digitális jel másik fő jellemzője, amit *kvantálási hossz*nak mondunk. A jel pontossága és a visszaalakításkor kapott eredmény nagymértékben függ a tartományok számától, így a kvantálási hosszról.

A következő lépés a kódolási művelet. Ebben a műveletben minden mintavételi pontban felvett jelhez a kvantálás végeredményének megfelelő bináris kódszót rendelünk hozzá. Ezzel az eljárással minden  $y(m)$  jelhez bináris kódszót rendeltünk, tehát végeredményként kaptunk egy digitálisan feldolgozható jelet.

### 3.3.1.2. Digitális–analóg átalakítás

A digitális jel analóg jellé történő alakításakor (Digital to Analog Conversion – DAC) először a vett bináris jelsorozatot kell helyreállítani, amennyiben az szükséges. A jelsorozat visszaállítását követően a kódszavakból előállítható az  $y'(m)$  függvény, melyet egy alul-áteresztő szűrővel szűrve megkapható az eredeti  $x(t)$  jelet közelítő  $x'(t)$  analóg jel.

A közelítés javítására alkalmazhatunk interpolációt, mely a két függvény hasonlóságát növeli és csökkenti a torzítást. Az analóg jel mintavételezett formájából történő visszaállítás feltétele a Nyquist elméletben megadott. A Nyquist elmélet azt mondja ki, hogy egy sávkorlátozott jel esetében annak  $F_c$  legnagyobb frekvenci-



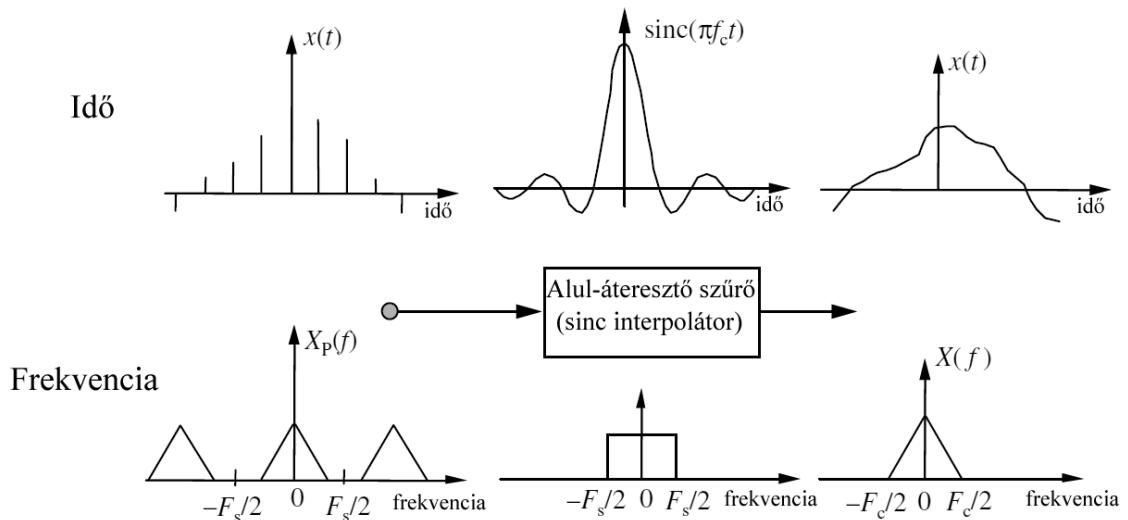
ájú összetevője is visszaállítható abban az esetben, ha a mintavételezés másodpercenkénti frekvenciája nagyobb volt a frekvencia kétszeresénél, vagyis  $F_s > 2F_c$ . Tegyük fel, hogy egy sávkorlátozott  $x(t)$  folytonos jel  $F_s$  másodpercenkénti mintavételezési frekvenciával mintavételezett. A diszkrét időközönként mintavételezett jel  $x(m)$  a következő szorzattal fejezhető ki:  $x(m) = x(t)p(t) = \sum_{m=-\infty}^{\infty} x(t)\delta(t - mT_s)$ , ahol

$p(t) = \delta(t - mT_s)$  a mintavételező függvény, és  $T_s = \frac{1}{F_s}$  a mintavételezési idő intervallum hossza. Vegyük az iménti egyenletet Fourier transzformálva, mely megmutatja, hogy a mintavételezett jel spektruma a következő egyenlettel írható fel:

$X_s(f) = X(f) * P(f) = \sum_{k=-\infty}^{\infty} X(f + kF_s)$ , ahol  $X(f)$  az  $x(t)$  jel spektruma, és  $P(f)$  a  $p(t)$

mintavételező függvénye, a  $*$  pedig a konvolúciós szorzás műveletét jelöli. Az iménti egyenlet állítása szerint a mintavételezett jel spektruma az alapsáv  $X(f)$  spektrumából és annak egyenletes  $T_s = \frac{1}{F_s}$  közönkénti ismétléséből vagy leképezéséből

áll össze. Ha a mintavételezési frekvencia felette van a Nyquist aránynak, akkor az alapsáv  $X(f)$  spektruma nincs átfedésben a  $X(f \pm kF_s)$  képeivel, és az eredeti jel



19. ábra Folytonos jel visszaállítása mintáiból. A frekvenciatartomány interpolációja az alul-áteresztő szűrővel ekvivalens

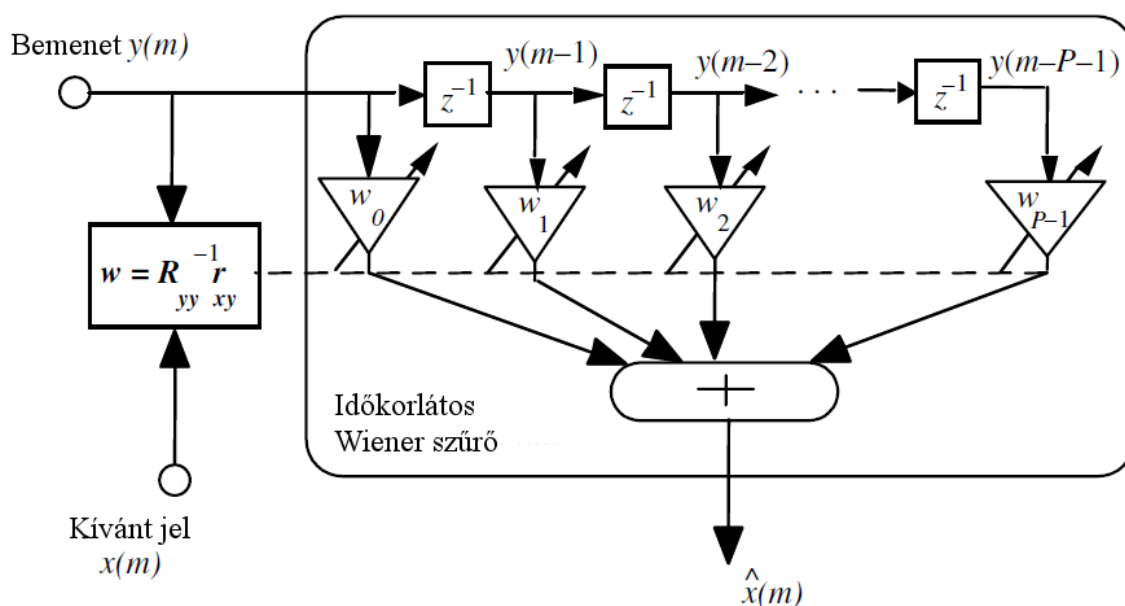
A gyakorlatban az  $F_s$  mintavételezési frekvencia nagyobb mint  $2F_s$ , általában  $2,5F_s$ , annak érdekében, hogy az átmeneti sáv szélességet hozzáigazítsuk az interpoláció alul-áteresztő szűrőjéhez.

### 3.3.2. A Wiener szűrő

A Wiener elméletet kidolgozójáról, Norbert Wiener-ről nevezték el, aki megfogalmazta az adatfüggő lineáris legkisebb négyzetek módszerére épülő szűrő elméletét. A Wiener szűrő központi szerepet játszik alkalmazásainak széles körében, mint például a visszhang érvénytelenítés, jelvisszaállítás, csatorna kompenzáció és rendszerazonosítás. A szűrő együtthatóit úgy kell kiszámítani, hogy a szűrő kimenő jele a lehető legjobban hasonlítson – vagyis különbségük négyzetének várható értéke legyen minimális – egy ismert úgynevezett megkívánt jelre. A Wiener elmélet feltételezi, hogy a jelek állandóak. Azonban ha a szűrő együtthatóit időnként újra kiszámítjuk, minden  $N$  elemű jelblokk után, a szűrő alakítja önmagát, a blokkokon belüli átlagos jellemzőkhöz, így blokk-alkalmazkodóvá válik. A blokk-alkalmazkodó (vagy szegmens-alkalmazkodó) szűrő alkalmazható olyan jelek esetében, mint például az emberi beszéd.

#### 3.3.2.1. A legkisebb négyzetek becslése

Wiener eredeti munkájában a folyamatos idejű legkisebb négyzetek módszerét adta meg, interpolációval. A módszert később kiterjesztette diszkrét időre is, mely egyszerűen megoldható, és a gyakorlatban jobban hasznosítható a digitális jelpro-



20. ábra Wiener szűrő felépítése

cesszorokban. A szűrő lehet végtelen hosszú jel feldolgozását célzó, vagy véges jel-hosszúságot figyelembevevő. Általában a végtelen szűrőnek egy halmaz nem lineáris egyenlet az eredménye, a véges hosszúságú jelre alkalmazott szűrőnek pedig egy halmaz lineáris egyenlet, mely egy zárt alakú megoldás. A véges szűrő viszonylag egyszerűbb számítást tesz lehetővé, működéséből adódóan stabilabb és praktikusabb is. Hátránya a végtelen szűrővel szemben, hogy a kívánt eredmény elérése érdekében nagyon sok együttható szükséges.

A 20. ábrán szemléltetett Wiener szűrő egy  $\mathbf{w}$  együttható vektorral megvalósított. A szűrő kap egy  $y(m)$  bemeneti jelet, és készít egy  $\hat{x}(m)$  kimeneti jelet, mely a legkisebb négyzetes eltérést mutatja az  $x(m)$  megkívánt vagy céljelhez képest. A szűrő bemenete és kimenete közötti összefüggés a következő egyenlettel írható fel:

$$\hat{x}(m) = \sum_{k=0}^{P-1} w_k y(m-k) = \mathbf{w}^T \mathbf{y}, \text{ ahol } m \text{ a diszkrét idő mérőszám, a szűrő bemeneti}$$

jele:  $\mathbf{y}^T = [y(m), y(m-1), \dots, y(m-P+1)]$ , és a paramétervektor (a Wiener szűrő együttható vektora)  $\mathbf{w}^T = [w_0, w_1, \dots, w_{P-1}]$ . A szűrési művelet kétféle egymással ekvivalens alakban fejthető ki, konvolúciós összegként, és a belső vektorok szorzataként. A hibajel  $e(m)$  a kimenőjel  $\hat{x}(m)$  és a megkívánt jel  $x(m)$  különbségeként számítható ki a következő összefüggéssel:  $e(m) = x(m) - \hat{x}(m) = x(m) - \mathbf{w}^T \mathbf{y}$ . Ebből az egyenletből kiderül, hogy egy adott  $y(m)$  bemenőjelre és  $x(m)$  megkívánt jelre a hibajel a  $\mathbf{w}$  együtthatóvektortól függ. A hibajel és az együtthatóvektor közötti összefüggés kiderítése érdekében kifejtve az előző egyenletet a következőt kapjuk:

$$\begin{pmatrix} e(0) \\ e(1) \\ e(2) \\ \vdots \\ e(N-1) \end{pmatrix} = \begin{pmatrix} x(0) \\ x(1) \\ x(2) \\ \vdots \\ x(N-1) \end{pmatrix} - \begin{pmatrix} y(0) & y(-1) & y(-2) & \dots & y(1-P) \\ y(1) & y(0) & y(-1) & \dots & y(2-P) \\ y(2) & y(1) & y(0) & \dots & y(3-P) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y(N-1) & y(N-2) & y(N-3) & \dots & y(N-P) \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_{P-1} \end{pmatrix}.$$

Ez az egyenlet  $N$  mintára,  $y(m)$  és  $x(m)$  jelekre lett felírva. Tömörített alakban a következőképpen írható fel:  $\mathbf{e} = \mathbf{x} - \mathbf{Y}\mathbf{w}$ , ahol  $\mathbf{e}$  a hibavektor,  $\mathbf{x}$  a kívánt jel vektor,  $\mathbf{Y}$  a bementi jel mátrix, és  $\mathbf{Y}\mathbf{w}$  szorzat pedig a  $\hat{x}$  kimenőjel vektor. Feltételezzük, hogy  $P$  kiinduló minta ismert, vagy hanem akkor 0 értékűnek tekintendő. Ha az előbb felírt mátrixegyenletben az  $N$  és a  $P$  értéke egyenlő, akkor egy négyzetes mátrixegyenletet kapunk, és ez az egyetlen megoldás, ahol a számítási hiba értéke nulla. Ha  $N < P$  a minták száma  $N$  elégtelen mennyiségű ahhoz, hogy elérjük az egyedülálló megoldást a szűrő együtthatóinál, ezért ebben az esetben végtelen sok megoldás létezik nulla számítási hibával és a mátrixegyenletet határozatlannak mond-

juk. A gyakorlatban a jelminták száma sokkal nagyobb, mint a szűrő hossza – vagyis  $N > P$  –, ebben az esetben a mátrixegyenletet túlhatározottnak mondjuk, és létezik egy egyedi megoldás, általában nulla számítási hibával. Mikor  $N > P$  a szűrő együtthatói úgy lesznek kiszámítva, hogy az átlagos hiba minimális legyen, mint például a hiba átlagos értékének abszolút értéke  $E[|e(m)|]$ , vagy a hiba négyzetének középértéke  $E[e^2(m)]$ , ahol  $E[.]$  a várható érték operátor. A hibafüggvény kiválasztása hatással van a megoldás optimális mivoltára, és a számítási komplexitására. A Wiener elméletben az objektív kritérium a legkisebb négyzetes hiba (LSE<sup>2</sup>) a szűrő kimeneti jele és a kívánt jel között. A legkisebb négyzetes hiba kritérium optimális a normális eloszlást követő jelek esetében. A korlátozott idő intervallumra alkalmazott legkisebb négyzetes hiba kritériumú Wiener szűrő egy lineáris, zárt formulához vezet. A Wiener szűrő együtthatói egy átlagos négyzetes hiba függvény  $E[e^2(m)]$  minimalizálásával eszközli ki, tekintettel a szűrő  $\mathbf{w}$  együttható vektorára. A számítási hiba négyzetének középértéke a következő egyenlettel adható meg:  $E[e^2(m)] = E[(x(m) - \mathbf{w}^T \mathbf{y})^2] = E[x^2(m)] - 2\mathbf{w}^T E[\mathbf{y}x(m)] + \mathbf{w}^T E[\mathbf{y}\mathbf{y}^T] \mathbf{w} = r_{xx}(0) - 2\mathbf{w}^T \mathbf{r}_{yx} + \mathbf{w}^T \mathbf{R}_{yy} \mathbf{w}$ , ahol  $\mathbf{R}_{yy} = E[\mathbf{y}(m)\mathbf{y}^T(m)]$  a bemenő jel autokorrelációs<sup>3</sup> mátrixa,  $\mathbf{r}_{xy} = E[x(m)\mathbf{y}(m)]$  pedig a bemeneti jel és a kívánt jel kereszt-kölcsönösségi vektora. Az előző kifejezést kifejtve a következő egyenletet kapjuk:

$$E[e^2(m)] = r_{xx}(0) - 2 \sum_{k=0}^{P-1} w_k r_{yx}(k) + \sum_{k=0}^{P-1} w_k \sum_{j=0}^{P-1} w_j r_{yy}(k-j), \text{ ahol } r_{yy}(k) \text{ és } r_{yx}(k) \text{ az } \mathbf{R}_{yy}$$

autokorrelációs mátrix és az  $\mathbf{r}_{xy}$  kereszt-kölcsönösségi vektor elemei külön-külön. Az előző egyenletből a korlátozott idő intervallumra alkalmazott szűrő négyzetes hibájának középérték számítására egy másodfokú függvény alkalmazható a  $\mathbf{w}$  együttható vektor értékeire, mely függvénynek egyetlen minimuma van. Például ha szűrőnek csak két együtthatója van ( $w_1, w_2$ ), akkor a négyzetes hiba középérték függvénye egy medence alakú képet mutat, egyetlen minimummal. A legkisebb négyzetes hiba jelenti az energiaminimumot.

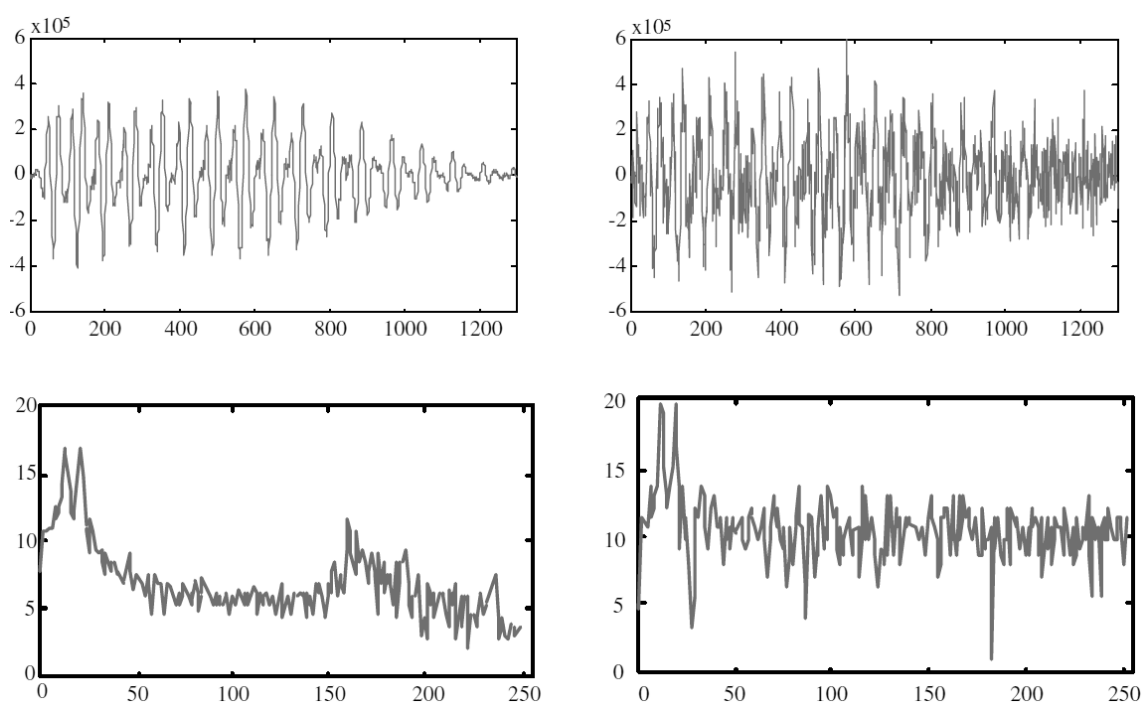
### 3.3.3. Zajcsökkentés spektrális kivonással

A spektrális kivonás módszerének lényege, hogy a zajos jelből veszünk egy olyan szakaszt, ahol a hasznos jel nincs jelen, tehát csak zaj észlelhető, és ezt a jelet egy megfelelő módszerrel kivonjuk a zajos jelből. Ez megtehető oly módon, hogy a jelek statisztikáit és spektrális paramétereit vesszük alapul. Ezzel a módszerrel adott esetben jelvisszaállítás is lehetséges.

A problémára a legegyszerűbb megoldást jelentené, ha külön rendelkezésre állna a zajos jel és a zaj. Ebben az esetben egy egyszerű művelettel eltávolítható lenne a zaj. A zajérvénytelenítő, melynek működését már korábban kifejtettem pontosan ilyen elven működik. Általában azonban csak a zajos jel áll rendelkezésre, tehát nincs lehetőség a zaj érvénytelenítésére, de előfordulhat, hogy a zaj hatását csökkenteni tudjuk a hasznos jel spektrumán.

### 3.3.3.1. A spektrális kivonás módszere

A 21. ábra olyan jeleket mutat, melyek zajjal terhelték, de nincs lehetőség a zaj érvénytelenítésére. A hasznos jel spektrumának középvértéke növelhető úgy, hogy a

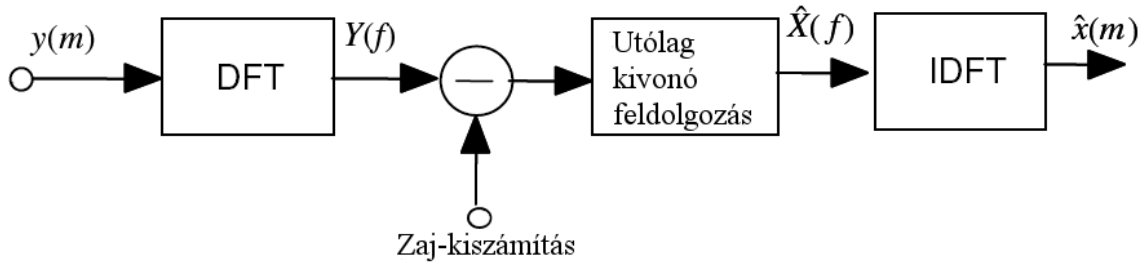


21. ábra Zajjal terhelt jelek alakjai az idő függvényében

zajos jel spektrumának középvértékéből kivonjuk a zaj spektrumának középvértékét. A zajos jel időben a következőképpen modellezhető:  $y(m)=x(m)+n(m)$ , ahol  $y(m)$  jelöli a zajjal terhelt jelet,  $x(m)$  a hasznos jelet,  $n(m)$  a zajt,  $m$  pedig a diszkrét idő mérőszám. Egy frekvenciatartományban az előző egyenlet a következőképpen fejthető ki:  $Y(f)=X(f)+N(f)$ , ahol  $Y(f)$  az  $y(m)$  zajos jel Fourier transzformáltja,  $X(f)$  az  $x(m)$  hasznos jel Fourier transzformáltja,  $N(f)$  az  $n(m)$  zaj Fourier transzformáltja, és  $f$  a frekvenciaváltozó. A spektrális kivonásban az  $y(m)$  zajos jelet vesszük, ezt egy átmeneti tárolóba helyezzük, majd felosztjuk  $N$  egyenlő részre. Minden szegmens esetében Hann vagy Hamming ablak segítségével szűrjük a spektrumot, ezután transzformáljuk diszkrét Fourier transzformációval (DFT) az összes  $N$  darab spektrális mintát. Az ablakok minden szegmens végén csökkentik annak hatását,

hogy a vett jelet feldaraboltuk, vagyis nem folytonos jeleket illesztünk össze. Az ablakozott jel így adható meg:  $y_w(m) = w(m)y(m) = w(m)[x(m) + n(m)] = x_w(m) + n_w(m)$ . Az alakozási eljárás a frekvenciatartományokra az alábbi módon fejthető ki (DFT segítségével):  $Y_w(f) = W(f) * Y(f) = X_w(f) + N_w(f)$ , ahol a  $*$  operátor a konvolúciós szorzást jelenti.

A következő ábra a spektrális kivonás módszerének blokkdiagramját szemlélteti. (Az IDFT jelölés a diszkrét Fourier transzformáció inverzét jelenti)



22. ábra A spektrális kivonás egyszerűsített blokkdiagramja

A spektrális kivonás egyenlettel a következő módon írható le:  $|\hat{X}(f)|^b = |Y(f)|^b - \alpha \overline{|N(f)|^b}$ , ahol  $|\hat{X}(f)|^b$  az eredeti jel spektrumának kiszámítása,  $|X(f)|^b$  és  $|N(f)|^b$  pedig az időátlagolt zaj spektrum. A spektrális kivonás nagyságához a  $b$  kitevő értékét 1-re, az erősségéhez pedig 2-re kell választani. Az  $\alpha$  paraméter a kivont zaj mennyiségét szabályozza. A teljes kivonáshoz  $\alpha=1$  értéket kell választani, a túlvonáshoz pedig  $\alpha>1$  érték lesz a megfelelő. Az időátlagolt zajspektrum beszerezhető a jelmentes szakaszokból, ahol csak kizárólag a zaj van jelen, nincs hasznos jel, és az alábbi összefüggés segítségével formalizálható:

$$\overline{|N(f)|^b} = \frac{1}{K} \sum_{i=0}^{K-1} |N_i(f)|^b, \text{ ahol } |N_i(f)|^b \text{ a spektrum } i\text{-edik zajkerete, és feltesszük,}$$

hogy a zaj  $K$  keretből áll, és  $K$  változó. Más esetben az átlagolt zajspektrum beszerezhető az elsőrendű digitális alul-áteresztő szűrő kimeneteként a következőképpen:

$$\overline{|N_i(f)|^b} = \rho \overline{|N_{i-1}(f)|^b} + (1 - \rho) |N_i(f)|^b, \text{ ahol az alul-áteresztő szűrő együtthatója}$$

$\rho$ , melynek tipikus értéke 0,85 és 0,99 közé szokott esni. Az időtartomány alapú jelvisszaállításhoz a nagyságspektrum  $|\hat{X}(f)|$  kiszámítását kombináljuk a zajos jel fázisával, majd ezt követően transzformáljuk időtartományba az inverz diszkrét Fourier transzformáció segítségével, mely a következőképpen írható fel:

$$\hat{x}(m) = \sum_{k=0}^{N-1} |\hat{X}(k)| e^{j\theta_Y(k)} e^{-j\frac{2\pi}{N}km}, \text{ ahol } \theta_Y(k) \text{ a zajos jel fázisa } Y(k) \text{ frekvencián. Ez a jel-}$$

viSSzaállító egyenlet feltételezi, hogy a hallható tartományba eső zaj főként a nagyságspektrum torzítása és a fázistorzítás miatt nagyrészt nem hallható. A szimulált fázistorzítás megfigyelhető hatásának becslése igazolta a feltevést. A zaj spektrumának változatossága következtében előfordulhat, hogy nagyság és az energiaspektrum szempontjából negatív eredményt kapunk a spektrális kivonással. Ennek végeredményeképpen nagyon valószínű, hogy csökkenni fog a jel-zaj arány. A negatív eredmények elkerülése érdekében a spektrális kivonás kimeneti jelét egy utófeldolgozó térképező függvénnyel feldolgozzuk, melyet a következőképpen formalizálhatunk:  $T[|\hat{X}(f)|] = \begin{cases} |\hat{X}(f)| & \text{ha } |\hat{X}(f)| > \beta |Y(f)| \\ \text{fn}[|Y(f)|] & \text{egyébként} \end{cases}$ . Például választhatunk egy

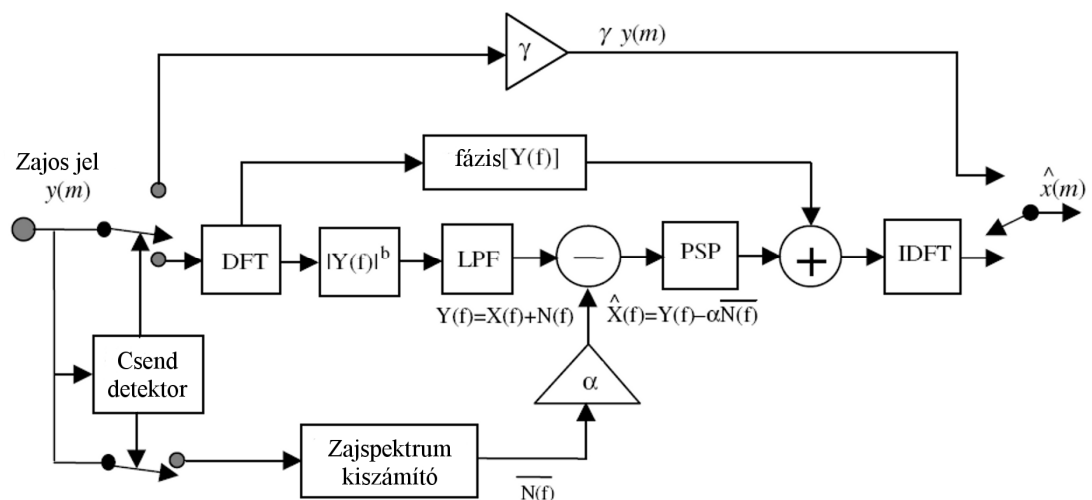
olyan szabályt, mely ha a  $|\hat{X}(f)| > 0,01 |Y(f)|$  (a nagyság spektrum szorzása 0,01-dal ekvivalens a  $-40\text{dB}$ -lel) érték igaz, akkor az  $|\hat{X}(f)|$  érték valamilyen függvénye legyen a zajos jelnek  $\text{fn}[|Y(f)|]$ . Ennek legegyszerűbb formája, hogy beállítunk egy  $\text{fn}[|Y(f)|] = \text{zajszint}$  értéket, ahol a zajszint pozitív konstans. Egy másik lehetőség, ha  $\text{fn}[|Y(f)|] = \beta |Y(f)|$ . Ez esetben:

$$T[|\hat{X}(f)|] = \begin{cases} |\hat{X}(f)| & \text{ha } |\hat{X}(f)| > \beta |Y(f)| \\ \beta |Y(f)| & \text{egyébként} \end{cases}$$

A spektrális kivonás energiaspektrum vagy nagyságspektrum tartományok alapján is implementálható. A két módszer hasonló, de elméletileg eredményeznek némi különbséget a várható teljesítményben.

### 3.3.3.2. A spektrális kivonó megvalósítása

Egy spektrális zajkivonó rendszer blokkdiagramját mutatja az alábbi ábra,



23. ábra Spektrális kivonó megvalósításának blokkdiagramja

mely a következő alrendszeréből áll:

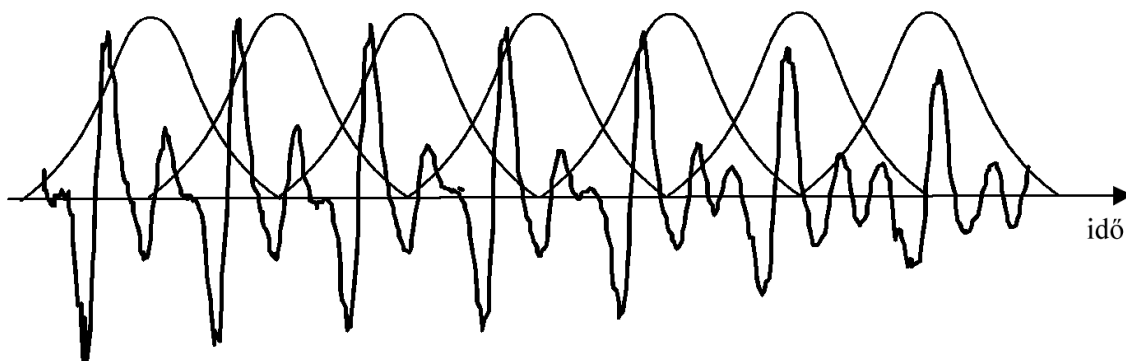
- a.) egy csend detektor a jelmentes időszakok megtalálására, a zaj spektruma ezen időtartamok alatt frissül.
- b.) a diszkrét Fourier transzformációt megvalósító rész (DFT) az időtartományt frekvenciatartományá alakítja, a DFT egységet egy nagyság operátor követi.
- c.) alul-áteresztő szűrő (Lowpass Filter, LPF), a zaj egyenetlenségeinek csökkentésére, melynek célja a zaj ingadozásaiból adódó feldolgozási torzítás csökkentése.
- d.) utófeldolgozó, a spektrális kivonás eredményeképpen keletkezett torzítások eltávolítására.
- e.) inverz diszkrét Fourier transzformációt végző rész, mely frekvenciatartományból viszi a feldolgozott jelet időtartományba.
- f.) csillapító berendezés  $\gamma$ , mely a zaj csillapítását végzi a csöndes részeken.

A diszkrét Fourier transzformáción alapuló spektrális kivonó blokkdiagramja egy algoritmus. A beérkező audiojelet átmeneti tárolóba rakjuk, majd  $N$  darab egymást átfedő mintablokkra osztjuk. Minden blokk Hann vagy Hamming ablakolt, és azt követően diszkrét Fourier transzformációval átalakítjuk frekvenciatartományá. A spektrális kivonás után a jel nagyságspektrumát kombináljuk a zajos jel fázisával, és visszaalakítjuk időtartományá. Minden blokkot egymáshoz illesztünk, úgy, hogy legyen átfedés a megelőző és a rákövetkező blokkal, és ez adja a végleges kimeneti jelalakot.

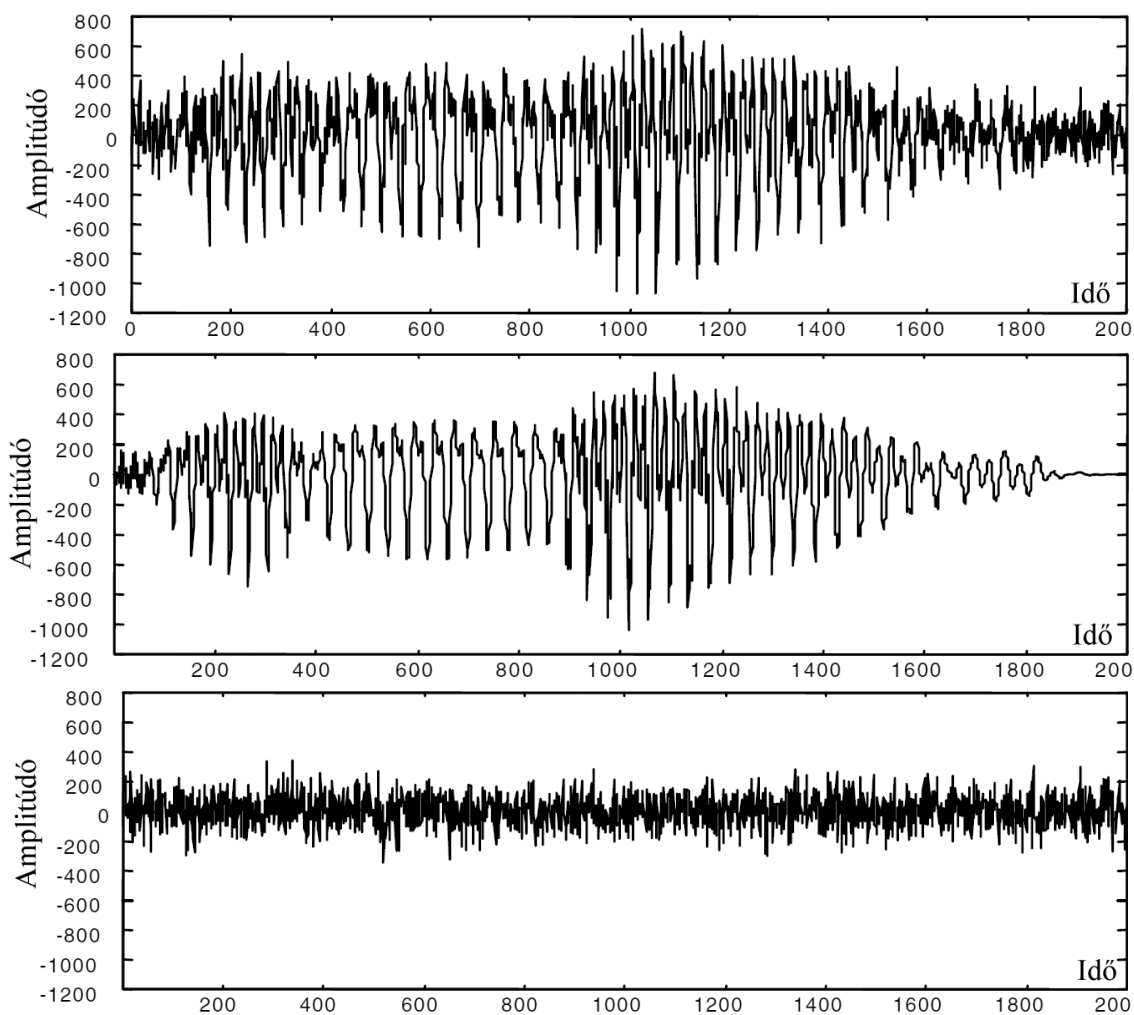
A blokk hosszának megválasztása a spektrális elemzéshez egy kompromiszsum, mely két ellentmondó szükséglet, az időfelbontás és a spektrális felbontás igényeit kell, hogy kielégítse. Tipikusan 5–50 ms közé esik a használt blokkok hossza. Ez egy 20 kHz-es mintavételezési frekvencián azt jelenti, hogy  $N$  értéke 100–1000 tartományba esik. A spektrum frekvenciafelbontása egyenesen arányos a minták számával ( $N$ ). Nagyobb  $N$  esetén a spektrum értékelése jobb lesz. Ez csak részben igaz a frekvenciatartomány alsó részére, mivel az alacsonyabb frekvenciájú részek az idő múlásával lassan változnak, és nagyobb ablak szükséges a stabil értékeléshez. Az ellentmondó szükséglet az audiojelek nem stacionárius természetéből adódnak, tehát az ablak méretének azért nem szabad túl nagyoknak lennie, hogy a rövid időtartamú események ne mosódjanak el.

Az ablakok átfedéseinek (24. ábra) fő feladata, hogy csillapítsák a blokkok végpontjain keletkező folytonossághiányokat, melyek a darabolásból adódnak. Bár



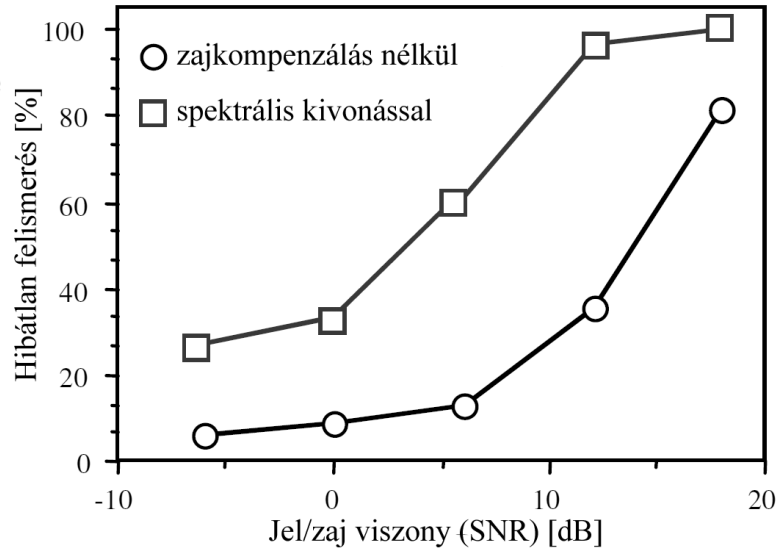


24. ábra A spektrális kivonásban alkalmazott ablakolási és átfedési folyamat létezik számos felhasználható ablaktípus, különböző karakterisztikákkal, mégis a legtöbb megvalósításban Hamming ablakokat használnak. A torzítás eltávolítás során – melyet a spektrális kivonás eredményezett – az utófeldolgozó algoritmus olyan információkat használ fel, mint a frekvenciacsatornák összefüggése az egymást követő blokkokban, valamint az jel eseményeinek és torzításának időtartamai.



25. ábra Zajos jel (felső), zajmentes jel (középső), és zaj (alsó)

A jel spektrális összetevőinek összefüggése – az időtengely mentén – részben szabályozható az ablakhossz megválasztásával és az átfedéssel. Ez az összefüggés növekszik az ablak hosszának csökkentésével és az átfedés növelésével. Az átfedés növelése azonban növeli a zajfrekvenciák összefüggését is, az időtengely mentén.



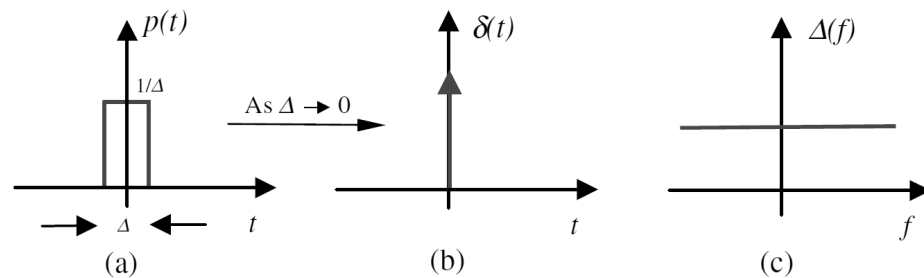
26. ábra A spektrális kivonás jelentősége a beszéd felismerésben

### 3.3.4. Az impulzív zaj eltávolítása

Az impulzív zajok gyors be/ki kapcsolásokhoz hasonló impulzusokat produkálnak és nagyon kiszámíthatatlan a megjelenésük, tehát bármely időpillanatban előfordulhatnak, és bármekkora időközrel. Ezen zajok eltávolítására és csökkentésére az előző részben kifejtett spektrális kivonás módszere nem alkalmas, mert itt a zajszint nem határozható meg.

#### 3.3.4.1. Impulzív zaj matematikai modellje

Az analóg impulzív zajt matematikai értelemben az alábbi ábra segítségével szemléltethetjük. Ha megvizsgáljuk az ábra (a) részén szemléltetett egységnyi te-



27. ábra (a) Egységnyi területű lökés, (b) impulzus, (c) az impulzus spektruma

rületű lökéshullám  $p(t)$  (négyzög jel) viselkedését, ahogy a lökés szélessége  $\Delta$  nullához tart, a lökés impulzussá válik. Az impulzus függvényt az ábra (b) része mutatja, és egy végtelenül kicsi időszélességgel rendelkező lökést a következőképpen

definiálhatunk:  $\delta(t) = \lim_{\Delta \rightarrow 0} p(t) = \begin{cases} \frac{1}{\Delta}, & |t| \leq \frac{\Delta}{2} \\ 0, & |t| > \frac{\Delta}{2} \end{cases}$ . Az impulzusfüggvény integrálja a kö-

vetkező:  $\int_{-\infty}^{\infty} \delta(t) dt = \Delta \times \frac{1}{\Delta} = 1$ . Az impulzusfüggvény Fourier transzformáltja felállítha-

tó:  $\Delta(f) = \int \delta(t) e^{-j2\pi ft} dt = e^0 = 1$ , ahol  $f$  a frekvenciaváltozó. Az impulzusfüggvényt

teszt jelleggel is szokták alkalmazni, a rendszer impulzusreakciójának megállapítására. Az ábra (c) részén mutatottak szerint az impulzus egy spektrálisan nagyon gazdag jel, minden frekvenciát azonos mennyiségben tartalmaz, tehát ezért alkalmas a tesztelésre.

Egy digitális impulzus  $\delta(m)$  (8. oldal 4. ábra (a) részén látható) úgy definiálható, mint egy egy minta hosszúságú „be” jel, és kifejezhető:  $\delta(m) = \begin{cases} 1, & m = 0 \\ 0, & m \neq 0 \end{cases}$ , ahol az

$m$  változó jelöli a diszkrét idő mérőszámot. A Fourier transzformációt használva

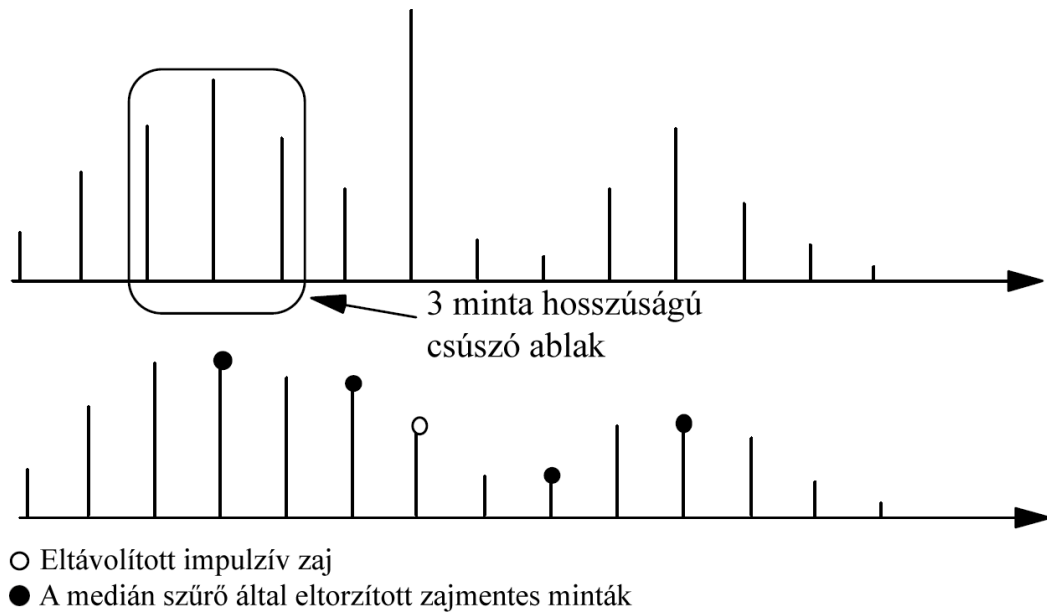
egy digitális impulzus frekvenciaspektruma megadható:  $\Delta(f) = \sum_{m=-\infty}^{\infty} \delta(m) e^{-j2\pi fm} = 1, 0,$

$-\infty < f < \infty$ .

Kommunikációs rendszerekbe a valódi impulzív típusú zajok hossza normálisan nagyobb, mint egy minta.

#### 3.3.4.2. A medián szűrő

Az impulzív zaj eltávolításának klasszikus megközelítése a medián szűrő. Medián alatt értjük az  $\{x(m)\}$  mintahalmaz elemeiből azt az  $x_{\text{med}}(m)$  értéket, melyre igaz, hogy a halmaz elemeinek fele ezen  $x_{\text{med}}(m)$  értéknél nagyobb, a másik fele pedig az  $x_{\text{med}}(m)$  értéknél kisebb. Tekintsük a mediánt minták egy halmazának, melyet úgy kapunk meg, hogy a mintákat rendezzük növekvő vagy csökkenő sorrendbe és kiválasztjuk a középső értéket. A medián szűrés esetén egy előre meghatározott hosszúságú ablakot csúsztatunk folyamatosan a jelen, és az ablakban található középső mintát kicseréljük az ablak mintáinak mediánjára, ahogy azt a következő ábra mutatja.



28. ábra Medián szűrő be- és kimenete. A működés közben torzítás is keletkezhet

A medián szűrő  $\hat{x}(m)$  kimenete az  $y(m)$  bemenet és egy  $2K+1$  mintát tartalmazó hosszúságú medián ablak alapján:

$$\hat{x}(m) = y_{\text{med}}(m) = \text{medián}[y(m-K), \dots, y(m), \dots, y(m+K)]$$

Egy számhalmaz mediánja a halmaz nem lineáris statisztikája egy nagyon jól hasznosítható tulajdonsággal, hogy nem érzékeny egy szokatlanul nagy mintaérték jelenlétére a halmazban, melyet kívülállónak hívunk. Ezzel szemben a középérték és részben a variancia érzékeny egy számhalmazban jelen levő kiugró értékre, vagyis az impulzív típusú zajra. Fontos tulajdonsága a medián szűrőnek – mely részben használható a képfeldolgozásban –, hogy megőrzi az éleket vagy a lépésenkénti folytonossághiányokat a jelben. A medián szűrők hatékonyan használhatók a képfeldolgozásban az impulzív zaj eltüntetésére az élek elmosása nélkül – ez nagyon jelentős a képfeldolgozásban. A medián filterrel történő kísérletezés során kiderült, hogy nem alkalmas magas minőségű eredmények produkálására audiojelek visszaállítása esetében. A medián szűrő nem tud mit kezdeni az olyan impulzív jellegű zajjal, mely hosszabb, mint egy vagy két minta. Továbbá a medián szűrő a hibásan impulzív zajnak vélt jelek módosításával jelentős mennyiségű feldolgozási torzítást okoz. A medián szűrő teljesítménye javítható egy alkalmazkodó határérték bevezetésével, tehát a mintát csak akkor cseréljük ki, ha annak eltérése a középső érték-

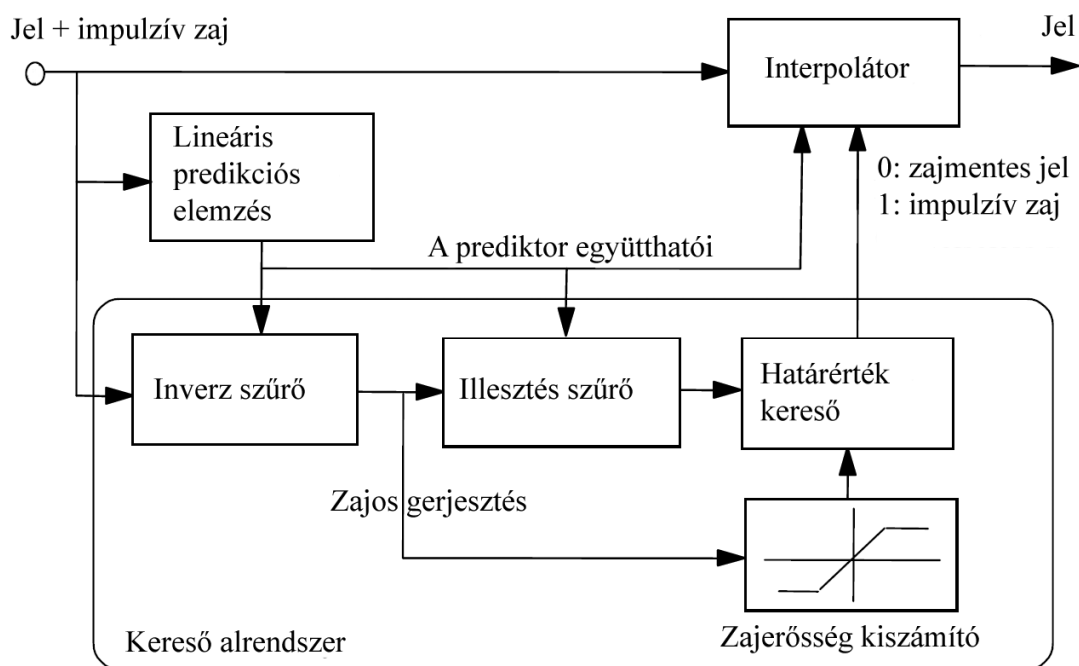
től meghaladja a határértéket:  $\hat{x}(m) = \begin{cases} y(m) & \text{ha } |y(m) - y_{\text{med}}(m)| < k\theta(m) \\ y_{\text{med}}(m) & \text{egyébként} \end{cases}$ , ahol

$\theta(m)$  egy alkalmazkodó határérték, mely erőteljes becsléssel határozza meg

$|y(m) - y_{\text{med}}(m)|$  átlagértékét,  $k$  pedig egy hangoló paraméter. A medián szűrők nem optimálisak, mert nem használhatók hatékonyan a jelgenerálás fiziológiájának előismerete vagy a jel és a zaj statisztikai eloszlásának modellezése nélkül.

### 3.3.4.3. Impulzív zaj eltávolítása lineáris prediktív modell segítségével

Az impulzív zavarok általában a jel egészét vizsgálva csupán csekély  $\alpha$  hányadát szennyezik. Mivel a jel jelentős  $1-\alpha$  része impulzív zajjal nem szennyezett, ezért előnyösebb megkeresni a zajos részeket (egyedülálló impulzusokat), és csak a torzult mintákat javítani. Ez a stratégia kiküszöböli a felesleges feldolgozásokat, és a minták egy viszonylag nagy része – mely nem szennyezett impulzív zajjal – érintetlenül eredeti állapotában marad. A következő ábra egy ilyen rendszer blokkdiagramját mutatja, mely két alrendszerből áll, egy keresőből és egy interpolátorból. A



29. ábra Egy impulzív zajeltávolító blokkdiagramja.

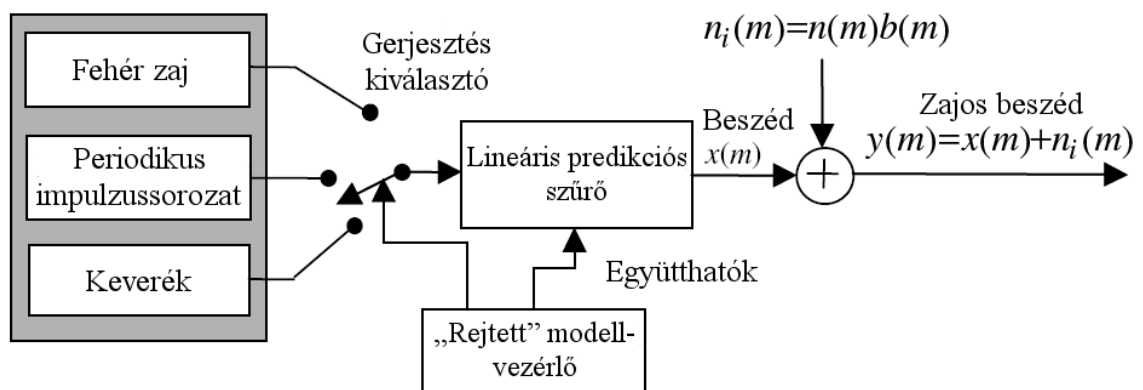
detektor megkeresi az impulzusok helyét, az interpolátor pedig kicseréli a torzult mintákat, figyelembe véve a torzult minta mindkét oldalán elhelyezkedő további mintákat. A detektor rész egy lineáris predikciós elemzőből, egy illeszkedésvizsgáló szűrőből és egy határérték keresőből áll. A kereső kimenete egy kétállású kapcsoló, mely az interpolátort vezérli. Impulzív zaj hiányában az interpolátor nincs bekapcsolva. Impulzív zaj jelenléte esetében az interpolátor bekapcsol és kicseréli a zaj által törölt mintákat.

## 3.3.4.3.1. Az impulzív zaj keresése

A legegyszerűbb módszer az ilyen jellegű zajok megtalálására, hogy meghatározzunk egy határértéket, és a határérték fölé eső mintákat zajnak osztályozzuk. Ez a módszer tökéletesen működik nagyon nagy impulzusok esetében, de ha a zaj amplitúdója a megadott határérték alá esik, akkor nem találja meg. A keresés eredményesebbé tehető, ha kihasználjuk a jel és az impulzív zaj jellemzőinek különbségét. Az impulzív zaj, vagy rövid időtartamú hullámvész bevisz egy a jelre nem jellemző megszakítást a jelbe. A megszakítás könnyebben megtalálható, ha a jelet szétválasztjuk. A szétválasztás (vagy digitális jelnél megkülönböztetés) művelete egyenértékű a kölcsönösségi viszony megszüntetésével, vagy a spektrális fehérítéssel. A lineáris prediktor modellezi a jel kölcsönösségi struktúráját, és ezen összefüggések megszüntetését végzi az inverz szűrő. A zajmentes  $x(m)$  jel lineáris predikciós modell segítségével a következő összefüggéssel írható fel:

$$x(m) = \sum_{k=1}^P a_k x(m-k) + e(m), \text{ ahol } a=[a_1, a_2, \dots, a_P]^T \text{ a lineáris prediktor } P \text{ rendű}$$

együttható vektora, és az  $e(m)$  gerjesztés egy zajhoz hasonló jel, vagy vegyes véletlenszerű zaj keveréke, és hullámvészok kvázi periodikus sorozata, melyet a következő ábra mutat. Az impulzív zajkereső nem a korrelálatlan kétállapotú impulzív zaj



30. ábra Zajos beszédmodell. A jelet egy lineáris prediktor modellezi, az impulzív zaj modellezésére pedig egy kétállapotú amplitúdómodulált modellt használunk.

alapján dolgozik, hanem azon a megfigyelésen alapszik, hogy a lineáris prediktorok jó modellt biztosítanak korrelált jelekre. A zajos  $y(m)$  jel a prediktor gerjesztett jelévé történő alakítása a következő hatásokat váltja ki:

- a.) Az jel amplitúdójának léptéke csökken az eredetileg gerjesztett jeléhez képest, mivel a zaj amplitúdójának léptéke változatlan marad, vagy növekszik.

- b.) A jel nem lesz kölcsönösen összefüggő, mert az impulzív zaj elmaszatólódott és átalakult léptékes alakba az inverz szűrő impulzus válaszaként.

Mindkét hatás javítja az impulzív zaj megtalálhatóságát. A beszéd és a zene hang időtartamának rezonanciája vagy a hangszer által erősített és spektrálisan alakított gerjesztésekből áll össze. Ez a gerjesztés véletlenszerűbb a beszédnél, és gyakran sokkal kisebb az amplitúdó-tartománya. Az impulzív zaj megtalálhatósága javítható az inverz szűréssel, mely lényeges lehet és függ a jel időben változó korrelációs struktúrájától. Ez a módszer egy optimális határértékképző használatával képes hatékonyan csökkenteni az impulzív zaj keresésének problémáját, a sorból kilógó minták megtalálását.

### 3.3.4.3.2. Az impulzív zaj megtalálhatóság-javításának elemzése

A következőkben az impulzív zaj megtalálhatóságának javítását elemezzük ez inverz szűrő kimenetének vizsgálatával. Az előző egyenlet használatával a zajos jel modelljét a következő egyenlet segítségével (az előző átírt változata) írhatjuk fel:

$$y(m) = x(m) + n_i(m) = \sum_{k=1}^P a_k x(m-k) + e(m) + n_i(m), \text{ ahol } x(m), y(m) \text{ és } n_i(m) \text{ rendre a}$$

zajos jel, a jel és a zaj. Egy számítást használva  $\hat{\mathbf{a}}$  lesz  $\mathbf{a}$  prediktor együttható vektora, és a zajos  $y(m)$  jel inverz szűrhető és átalakítható  $v(m)$  zajosan gerjesztett

$$\text{jellé így: } v(m) = y(m) - \sum_{k=1}^P \hat{a}_k y(m-k) = x(m) - n_i(m) - \sum_{k=1}^P (a_k - \tilde{a}_k)[x(m-k) + n_i(m-k)],$$

ahol  $\tilde{a}_k$  a prediktor együttható számítási hibája. A zajos jel modelljének egyenlete ez alapján az egyenlet alapján következő egyenleté írható át egy másik alakba:

$$v(m) = e(m) + n_i(m) + \sum_{k=1}^P \tilde{a}_k x(m-k) - \sum_{k=1}^P \hat{a}_k n_i(m-k). \text{ Ebből az egyenletből három}$$

tétel határozható meg, mely elősegíti a zaj megtalálását a gerjesztési szekvenciában.

- az  $n_i(m)$  impulzív zavar, mely általában domináns;
- a régebbi  $P$  darab impulzív zajminta hatása az inverz szűrő működése végett, a jelenbe maszátolva:  $\sum \hat{a}_k n_i(m-k)$ ;
- a gerjesztett jel varianciájának növekedése a paramétervektor számítási hibái miatt, mely a következő kifejezéssel írható le:  $\sum \tilde{a}_k x(m-k)$

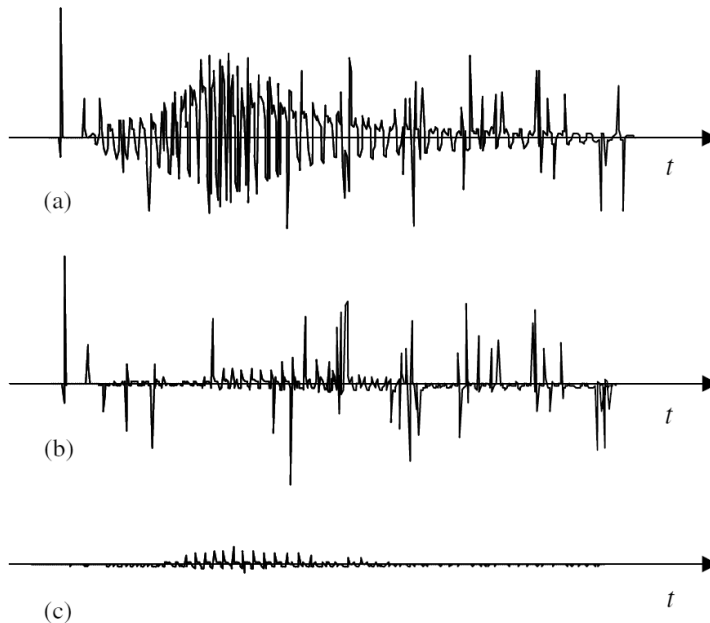
Az inverz szűrő eredményeinek javítása a következő formulákkal írható le. Az impulzív zaj a zajos jelhez viszonyított aránya az alábbiakban ismertetett módon

adható meg:  $\frac{\text{impulzív zaj energiája}}{\text{jel energiája}} = \frac{E[n_i^2(m)]}{E[x^2(m)]}$ , ahol  $E[\cdot]$  a várható érték operátora.

Az impulzív zaj detektálásában az érdeklődés tárgyát képező jel az impulzív zaj lesz, melyet a kísérő jelből kell megtalálni. Feltesszük, hogy a domináns zaj a  $v(m)$  zajos gerjesztési jelben az  $n_i(m)$  impulzív zaj, a két jel aránya a következőképpen írható fel:  $\frac{\text{impulzív zaj energiája}}{\text{gerjesztési energia}} = \frac{E[n_i^2(m)]}{E[e^2(m)]}$ . A teljes impulzív zaj jellel szembeni

aránya az előző két egyenlet hányadosából kapható meg, melyet a következő egyenlettel tudunk leírni:  $\frac{E[x^2(m)]}{E[e^2(m)]} = \text{nagyobbodás}$ .

Ez az egyszerű elemzés megmutatta, hogy az impulzív zaj megtalálhatósága függ a teljesítményerősítés karakterisztikájától, ennek következtében a lineáris predikciós modell rezonanciáitól. Beszéd esetében a jel amplitúdója zajmentes gerjesztett jelről lévén szó  $10^{-1}$  és  $10^{-4}$  között változik, ami a beszéd maga, ezért jelentős javulás érhető el az impulzív zaj megtalálhatóságában az inverz szűrő alkalmazásával a zajos beszédre. A következő ábra az impulzív zaj megtalálhatóságának



31. ábra Az inverz szűrő hatása a impulzív zaj megtalálhatóságára. (a) impulzív zajjal szennyezett beszéd, (b) az impulzív zajjal szennyezett beszéd gerjesztése, (c) beszéd gerjesztés impulzusmentesítve.

javítására alkalmazott inverz szűrő hatását mutatja. Az inverz szűrés hatása, hogy az  $x(m)$  jelet korrelálatlan  $e(m)$  jelgerjesztéssé alakítja, míg az impulzív zaj az inverz szűrő impulzív  $[1, -a_1, \dots, -a_p]$  zajátvitelének megfelelően torzul és megnövekszik, ahogy azt az iménti tétel mutatja a következő összefüggés alapján:

$\sum \hat{a}_k n_i(m-k)$ . Feltételezzük, hogy a gerjesztés egy

normális eloszlást követő fehér zaj, egy szűrő, melynek együtthatói az inverz szűrő együtthatóival egyeznek képes növelni a torzult impulzív zaj megtalálhatóságát a gerjesztési jelben.



## 3.3.4.3.3. Kétoldali prediktor az impulzív zaj megtalálására

Az impulzív zaj megtalálásának egyik hatékony módja, hogy a jelet korrelálatlanná tesszük. Ez a módszer felhasználható a kétoldali lineáris predikciós modellben is. Egy  $x(m)$  minta kétoldali lineáris predikciója  $P$  darab régebbi (múltbeli) és  $P$  darab jövőbeli minta alapján dolgozik, és a következő egyenlettel definiálható:

$$x(m) = \sum_{k=1}^P a_k x(m-k) + \sum_{k=1}^P a_{k+P} x(m+k) + e(m),$$

ahol  $a_k$  értékek a kétoldali prediktor együtthatói, és  $e(m)$  a gerjesztési jel. Minden elemzés arra az esetre készül, mikor csak egy oldalról vizsgáljuk a jelet, de ez kiterjeszthető a kétoldali modellre is. Azonban a gerjesztés mértéke a kétoldali modell esetében kisebb varianciát mutat, mint az egyoldali modellben, mivel az imént leírt egyenletben mind a múltbeli és mind a jövőbeli összes mintára modellezett a minták korrelációja. Habár az iménti egyenletben egy nem véletlenszerű szűrő van, az inverze szükséges a kereső alrendszerbe, mely már véletlenszerű. A kétoldali prediktor használata további javulást eredményezhet az impulzív zaj megtalálhatóságában.

## 3.3.4.3.4. Az eldobott minták interpolációja

Az impulzív zaj hatására a minták visszavonhatatlanul torzulnak, és eltávolításukkal „űr” keletkezik, melyet interpolációval próbálhatunk kitölteni. A hiányosságok interpolálására – ahhoz, hogy ne maradjon hallható nyomuk –, magas hanghűségű interpolátor szükséges. A legkisebb négyzetes autoregresszív (LSAR) módszer magas minőséget eredményez viszonylag kis számú az impulzív zaj miatt elveszett minta esetében. Az LSAR interpolációs módszer egy kétszintű folyamat. Az első szinten az impulzív zaj mindkét oldalán megmaradt minták alapján becsléssel meghatározzuk a jel predikciós modelljének paramétereit. A második szinten a meghatározott modellparaméterek és a hézag két oldalán található minták alapján interpoláljuk az elveszett mintákat. Ezzel az interpolációs módszerrel az impulzív zaj által eltorzított audiojelek magas minőségben visszaállíthatók.

## 4. Zajsűrés a gyakorlatban

### 4.1. Zenekari stúdiómunka

#### 4.1.1. A felvétel

Egy hangstúdióban legyen szó akár könnyűzenei vagy komolyzenei műről a hangrögzítés során számos problémával kell szembenéznünk. Amennyiben megoldható, az akusztikus (nem elektromos) hangszerek felvételét úgynevezett süket-szobában célszerű kivitelezni, a visszhanghatások elkerülése végett, melyek a hanganyag torzulását okozhatják. A felvételt adott esetben csatornákra bontva külön-külön rögzítjük. Ez azt jelenti, hogy az egyes hangszerek külön játsszák a darabot, majd a felvételt követően egy keverőpult segítségével születik meg a végleges hanganyag.

A felvétel során is felütik a fejüket a zajok. Amivel mindenképpen számolni kell, hogy a bemikrofonozott jelforrás valamilyen távolságra helyezkedik el a mikrofontól, néha a zenész átéléssel játszik, és a hangszerrel együtt mozog. A távolság és az alkalmazott mikrofon minőségének függvényében a rögzítő technika mindenképpen rögzít egy kevés környezeti zajt is. Ezen zajok ellen a környezet megfelelő megválasztásával lehet a legkönnyebben védekezni.

További problémákat okoznak a mikrofonok viszonylag alacsony jelszintjei, és a jeleket továbbító kábelek. A kábeleket célszerű lehetőség szerint minél vastagabbra, tiszta anyagból készültre, és minél rövidebbre választani, a lehető legkevesebb csatlakozóponttal, ugyanis ez tudja a jel legjobb minőségű szállítását elvégezni. A nagy keresztmetszet és a viszonylag kis hossz biztosítja a kis ellenállást, mely a jelveszteséget csökkenti, és a tiszta anyag (például oxigénmentes réz, négy kilences tisztasággal – 99,99 % réz) pedig a szennyezések miatt kialakuló örvényáramok okozta zajtól óvnak bennünket. A csatlakozópontok további torzítást és zajt eredményezhetnek, az anyagok különbözősége, és a csatlakozó felületek mérete miatt. További eredményjavulás várható az arany, vagy felületén aranyozott kábelek alkalmazásától a magas frekvenciájú jelek átvitele szempontjából – a skin effektus miatt. A skin effektus azt jelenti, hogy a váltakozó áramot szállító vezetékben az elemi töltéshordozók a vezetőben annál távolabb haladnak a vezető középpontjától, minél nagyobb az átvitt jel frekvenciája.

Az alacsony jelszint miatt a digitális feldolgozhatóság elérése érdekében előerősítő fokozatokat kell a rögzítő készülék és a jelforrás(ok) közé építeni, melyek esetében az alkatrészek kialakítástól és minőségétől függően termikus zajgenerátorként működnek. A sávátviteli problémákat többnyire nem az erősítőelemek okozzák, hanem a mikrofonok.

További érdekesség a jelforrás szempontjából, hogy az elektromos hangszereket, mint például egy elektromos gitár, vagy basszusgitár általában egy úgynevezett alap erősítőre csatlakoztatjuk, majd ezen erősítő hangszórója elé helyezünk el mikrofont. Általában e hangszerek esetében az effekteket már az erősítő előtt, vagy annak effektútjába csatlakoztatjuk. Ilyenkor egy torzító effekt képes a hangszer hangszedője (pickupja) által felfogott zajokat nagymértékben felerősíteni, amit még – egy a torzító előtt elhelyezett (a játéktechnikát könnyítő) – dinamika kompresszorral lehet „növelni”. Ezek a zajok elektromágneses zajok, és sajnos a hangszer működéséből adódóan nehéz ellenük védekezni, mivel a húrok rezgését is ilyen módon érzékeli a hangszedő. Ilyen hangszerek esetén már két termikus zajgenerátor rendszer kapcsolódik a rögzítő elé, ezzel fokozva a zajszennyezést.

A rögzítő technika manapság már többnyire digitális rendszer, melynek az érkező jeleket egy ADC segítségével digitálissá kell alakítania. A jel ekkorra már zajjal erősen szennyezett lesz, a korábbi berendezések felépítésének és minőségének függvényében. A digitalizálás a mai korszerű hangkártyáknak nem jelentenek problémát, mégis stúdió célokra külön gyártanak – magasabb árfekvésű, jobb minőségű – kártyákat. A mintavételezési frekvencia megválasztása ekkor az egyik fő probléma. Általában a végeredmény CD lemezre kerül, ahol szabvány írja elő a 44100 Hz-es mintavételezési frekvenciát. Vannak hangkártyák, melyek ilyen mintavételezési frekvencia mellett nem produkálnak lineáris sávátvitelt, ellenben 48000 Hz-en igen. A felvett jelet visszahallgatva, van különbség 48000 Hz-en és a 44100 Hz-en mintavételezett jelek között az első javára, bár azt gondolhatnánk, hogy a hallható tartományon már ezek a frekvenciák kívül esnek. Másik érdekes kérdés a kvantálási hossz. A mai világban már a megfelelő dinamika elérése érdekében 16 bitnél nagyobb értéket választunk.

#### 4.1.2. Az utómunka

Az utómunkálatok első lépése a felvétel során a jelre került zajok eltávolítása. Ez esetben, ha sávonként rögzítettünk, akkor minden sávra külön célszerű a zajt eltávolítani, egészen pontosan annak hatását csökkenteni, ugyanis a zajok vélet-

lenszerű mivolta nem teszi lehetővé teljes eltávolításukat. Az ilyen felvételek esetén azért kell külön kezelni a zajokat, mert egy mikrofon és egy elektromágneses hangszedő több zajgenerátorral együttműködve más és más jellegű zajokat eredményeznek. Ha például egy dobfelszerelésre gondolunk, akkor észrevehető, hogy a megfelelő minőségű jel érdekében célszerű a hangszer egyes darabjait külön be-mikrofonozni, adott esetben más típusú mikrofonokkal, tehát ekkor már egy mikrofonrendszerrel beszélünk. Az egyes mikrofonok hangját külön kell bontani hangszín szempontjából, esetleg sztereó felvétel esetében még a panoráma sem fog egyezni minden esetben (például: lábcintányér jobboldal, pergődob közép, kísérő cintányér baloldal, természetesen a zenész kezességét figyelembe véve lehet fordítva is). A hangszínt minden esetben az egyes részeken külön kell beállítani, ami függ a mikrofon minőségétől, távolságától, helyzetétől és állásától. Ez a mikrofonrendszer mikrofononkénti hangszín beállításokkal másfajta zajt állít elő, mint például az imént említett elektromos gitár hangszedője egy torzítóval. Ezért is fontos a sávonkénti felvétel. Ezek a zajok csak ritka esetben tartalmaznak impulzív jellegű zajokat (vagy ha az ilyen zajok megjelennek, azok megszüntetéséről még a felvétel előtt célszerű gondoskodni), ezért a zaj csökkentésére a legalkalmasabb módszer a spektrális kivonás. Fontos, hogy a zajszűrő program (algoritmus) olyan legyen, hogy a hangképet ne, vagy csak nagyon kis mértékben módosítsa, a legjobb eredmény elérése érdekében.

Ezt követően az egyes hangszerek még további effektek hozzáadásával színesíthetők, valamint itt nyílik lehetőség az utolsó hangszínbeállításra, ugyanis a felvételkor még csak nyersanyag keletkezik. Itt adott esetben az egyes sávokat további analóg keverőpultba küldhetjük, de léteznek olyan stúdióban alkalmazható programok, melyek képesek a többsávós felvétel kezelésére. Ilyen program például a korábbi (a 2003-ban az Adobe cég által megvásárolt Syntrillium cég fejlesztésében megjelent) Cool Edit, és a manapság helyette alkalmazott Adobe Audition program, melyek kereskedelmi forgalomban kaphatók, de van ingyenes időkorlátos próbaváltozatuk is. Ezek a programok nem csak többsávós hanganyag effektezésére, hangszínbeállítására alkalmasak, hanem az azt megelőző zajcsökkentésekre is.

Természetesen a megfelelő hangzás elérése érdekében a hanganyagot a hangfelvétel elkészülte után hetekkel célszerű keverni, majd a végeredményt többféle berendezésen visszahallgatni, a maximális minőség elérése érdekében. A megfelelő hangzás elérése után még egy dinamika kompresszor segítségével lehet az anyagot

„összetömöríteni”. A mintavételezési frekvencia és a kvantálási hossz a későbbi hanghordozó függvényében változhat a rögzítési értékekhez képest, de ez sajnos a jel torzulását és néha még további zajok megjelenését is okozhatja.

## **4.2. Hagyományos hanglemez, hangkazetta archiválása CD-re**

Minden archiválási folyamat az adott hanghordozón tártolt hanganyag digitalizálásával kezdődik. Ezt akár otthonunkban is megtehetjük, feltéve, hogy rendelkezünk a forrás hanghordozó megfelelő minőségű lejátszására alkalmas berendezéssel. A hangminőséget jelentősen befolyásolhatja a jelforrás, továbbá az is, hogy milyen eszközzel végezzük a digitalizálást, és hogy milyen kábelek kötik össze az adott berendezéseket. Itt is lehet környezeti elektromágneses zavarokra számítani.

### *4.2.1. Hanglemez archiválása*

#### 4.2.1.1. A lejátszó rendszer

A gyártók általában többféle megoldást alkalmaznak a lemezejátszó mechanikájának és a mechanikai mozgásokat elektromos jellé alakító részegységek kivitelezésére.

A mechanika leglényegesebb elemei:

- Lemeztányér, amely a lemez forgatásáért, megtartásáért felelős, továbbá sok esetben a hangkar mozgatásával is foglalkozik, az alsó mechanikus szerkezeten keresztül. Többnyire nagy tömegű, az egyenletes forgás megvalósítása érdekében (nagyobb tehetetlenség, így nem képes a motor gyors fordulatszám-változásainak követésére). Fontos a jó csapágyazás és kenés a könnyű futás érdekében.
- Motor, amely a lemeztányért hajtja, nagyon fontos, hogy állandó fordulatszámú legyen nagyon kicsi ingadozással, mely valamilyen általában erőzárás közlőművel kapcsolódik a lemeztányérhoz (ez esetben fordulatszáma sokkal nagyobb, mint a tányéré).
- Hangkar (esetleg a mozgató mechanikával együtt), mely hangszedő fejjel együtt a lemeztűt egy shellnek nevezett tartóban rögzíti, és gondoskodik a megfelelő tűnyomásról és a barázdakövetésről.

A lemeztányér meghajtására háromféle megoldást szoktak alkalmazni, az egyik a dörzshajtás, a másik a szíjhajtás – gyakorlatilag ez az egyik legelterjedtebb –, és végül a direkt hajtás – ez közvetlen a motor forgórészére rögzített viszonylag

könnyű lemeztányért jelent. A motor lehet váltakozó áramú, mely többnyire a szabványos hálózatban alkalmazott frekvenciára épít a forgás tekintetében, vagyis szinkron motor, és lehet úgynevezett DC (egyenáramú) motor is. A közvetlen hajtású lemezjátszók esetében a legpontosabb fordulatszám elérése végett kvarcvezérlést alkalmaznak. Egyes típusoknál egy stroboszkóp lámpa segíti a fordulatszám beállítását, egy potenciométer beiktatásával, melyet általában az egyenáramú motorral szerelt típusok esetében alkalmaznak. A három megoldás közül a kvarcvezérlésű direkt hajtás a legpontosabb megoldás.

A hangkart mechanikusan áthelyezni képes lemezjátszók a lemeztányér alatt elhelyezett mechanikus szerkezet segítségével teszik mindezt. Létezik belőlük fél-automata és automata kivitel is. A hangkarban helyet foglaló (sok esetben leszerelhető, cserélhető) shell biztosítja a lemeztű számára kialakított hangszedő fej rögzítését. A szögkiegyenlítés végett – nem gyakran ugyan – alkalmaznak úgynevezett tangenciális hangkart, melynél nincs ilyen jellegű probléma. A nem tangenciális esetben a fej valamekkora szöget zár be a hangkarral, melyet vagy a hangkar meghajlításával, vagy a shell kialakításával érnek el.

Hangszedőből háromféle alaptípus létezik, melyek működésükben különböznek egymástól.

Az egyik – régebben viszonylag olcsó és hatékony megoldást biztosító – típus a kristály hangszedő. Ez működését tekintve a gyémánt vagy zafír tű által a lemeztől letapogatott rezgéseket egy vékony cső segítségével egy piezokristályra vezeti, mely a rezgést elektromos jellé alakítja. Ez a letapogató rendszer viszonylag nagy túnyomással működik jól, és ez sajnos nem kíméli a mikrobarázdás lemezünket. Hangzás tekintetében nem túl jó, de kielégítő eredmény elérhető vele. Nagy előnye, hogy nem igényel külön korrekciós előerősítő fokozatot, viszont impedancia-illesztés szempontjából nagyimpedanciás bemenetet igen.

A másik szinte a legelterjedtebb, már jó hangzást biztosító megoldás a mozgó mágneses hangszedő. Ez esetben (sztereó kialakítást feltételezve) két egymással 90°-os szöget bezáró tekerccs között mozgatja a vékony cső végére rögzített mágnest a lemeztől letapogatott mechanikai rezgés. Túnyomás tekintetében változatos megoldásokkal találkozhatunk, mert van itt  $\frac{3}{4}$  grammtól kezdődően kristály hangszedőt megszegyenítő 5 grammos nyomást igénylő kialakítás is. Itt már kell egy – a köznapi életben alkalmazott erősítők nagytöbbségében az alkalmazás idő intervaluma alatt fellelhető – korrekciós előerősítő fokozat, mely a mély hangokat kiemeli, a közép hangokkal nem tesz semmit, és a magas hangokat pedig csillapítja. Erre

azért van szükség, mert a lemez előállításakor így rögzítik a jelet – a barázda szélességének csökkentése miatt (több hanganyag fér el egy lemezen). Viszonylag egyszerűen cserélhető az elkopott tű, melyet még manapság is forgalmaznak bizonyos helyeken.

A harmadik megoldás a mozgó tekercses megoldás. Ez csak fejjel együtt cserélhető, nagyon drága, és hangzásvilágát tekintve ez a legjobb minőségű. Az előerősítő fokozat ide is kell, de az erősítésének nagyobbak kell lennie az előzónél, mert a jel feszültsége csupán kb. az egytizede a mozgó mágneses hangszedőének. Létezik olyan erősítő, amely mindkét jelet képes fogadni. A megoldások közül a legjobb hangminőség a mozgó tekercses hangszedővel érhető el, de ára, és a szükséges további eszközök miatt inkább a mozgó mágneses hangszedőt használjuk.

#### 4.2.1.2. Várható zajok

A hanglemezek mechanikus adathordozók, ezért nagyon sérülékenyek. A nem megfelelő tárolás következtében a por komoly károkat tud okozni a hangzásban. A nem megfelelő poreltávolítás apró karcokat okoz, melyek hallhatók lesznek lejátszáskor. A lemez felületére az ujjainkról került zsír is károsítja a hanghordozót, továbbá a felületen nagyobb karcok is kialakulhatnak (a nem megfelelően mozgatott hangkar következtében a tű is okozhat ilyen karcokat), melyek az impulzív zajokért felelősek.

A hanglemez kialakítása során a felületi egyenetlenségek keletkeznek, melyek szintén hallhatóvá válhatnak lejátszáskor. Ezért van különbség a keleti és a nyugati nyomású lemezek között, mivel más a gyártás alapanyaga és némileg a technológiája is.

A lemezjátszó mechanikája is képes zajokat előállítani. A lemeztányér a csapágyazás és kenés függvényében valamilyen mértékben rezeg. Ez egy precízen csapágyazott motor esetén egy direkt hajtású lemezjátszóban nem okoz problémát, mert általában a hallható tartományon kívül esik a rezgés frekvenciája.

Különbséget kell tenni a lemez száraz, illetve nedves lejátszása között is. A nedves lejátszás esetén a lemezt mosófolyadékkal lejátszás közben mossuk, így a szennyeződések eltávolítása is megtörténik, továbbá a folyadék vékony réteget képez a tű és a barázda információt tartalmazó részei között, ezért a hangzás is változhat nagyon kis mértékben.

A lemez kivitelezésénél fogva a korrekciós előerősítő fokozat is zajt termel, majd az azt követő illesztő fokozat viselkedik zajgenerátorként. Ezekben az erősí-

tőkben az aktív elemek melegszenek, termikus zajt előállítva. Továbbá a nagy erősítés és a mágneses működés miatt érzékenyek az elektromágneses zajokra is, mely megfelelő szigeteléssel orvosolható.

Tehát a lemezen megjelenik a helyreállítandó hézagokat maga után vonó impulzív zaj, és a spektrumát tekintve nemfehér zaj is, melyeket két különböző eljárással tudunk csökkenteni.

#### 4.2.1.3. Zajcsökkentési eljárás

A felvétel elkészítésekor figyelni kell az alkalmazott technikák jelszint és impedancia-illesztésére, valamint a mintavételezési frekvenciára és a kvantálási hossza is. Lehetőleg a cél adathordozóéval kompatibilis értékeket kell választani, a minőségromlás elkerülése érdekében.

A lemezjátszók készülhettek a keleti, illetve nyugati régióban is. A keleti régió általában az ötpólusú tuchel aljzatokat alkalmazta a jelek továbbítására, amely nem csak kialakításában, hanem jelszintjében és impedanciájában is eltér a nyugati régióban elterjedt (4 darab) RCA csatlakozókhöz képest. A tuchel csatlakozók kisebb jelszinten dolgoznak, és nagyobb bemeneti impedanciát igényelnek, továbbá kisebb felületen érintkeznek, tehát nem alkalmasak közvetlenül digitalizálásra. (Természetesen egy egyszerű erősítőfokozat közbeiktatásával alkalmassá tehetők a feladat elvégzésére az ilyen berendezések is.) A célszerű választás tehát a folyamat kivitelezésére egy *nyugati gyártású lemezjátszó, mozgó mágneses* hangszedővel, a megfelelő előerősítő fokozatot tartalmazó szintén *nyugati gyártású erősítővel*, mely illeszthető a szabványos hangkártyák vonalbemenetére.

A folyamat a digitalizálással kezdődik. A feladatra a számítógépbe célszerű egy eleve jó értékű jel–zaj aránnyal rendelkező, kis torzítású hangkártyát választani. A Creative cég régebben is készített ilyeneket, és ma is gyárt alkalmas eszközöket. (A Sound Blaster sorozat elemei.) A digitalizáláskor a jel erős zajszenyezéssel érkezik az eszközbe.

Az első lépés az impulzív zajok megkeresése és eltávolítása. Erre alkalmas az előző részben említett Adobe Audition program. Itt a határértékek feltérképezése után indulhat a folyamat, sajnos az eredmény csak a végrehajtás után lesz hallgatható, nincs „előzetes bemutatás” lehetőség. A folyamat az impulzív zaj mennyiségétől és a számítógép sebességétől függően akár több órát is igénybe vehet. Közben látható, hogy hány mintát vélt zajnak, és hány mintát sikerült kijavítania a programnak. Természetesen az javított/kihagyott arány a beállítás függvénye, na-



gyon karcos lemez esetén sok impulzív zaj van, melyek javítása eredményezheti a hangminőség romlását is, oly formában, hogy a magas frekvenciájú jeleket is impulzív zajnak tekinti program.

A következő lépés a spektrális kivonás. Ekkor mintát kell venni egy csak zajos részből, mely lehet a lemez bevezető része. A mintában szereplő pillanatképek száma beállítható, de nem feltétlenül eredményez jobb minőséget a nagyobb szám. Fontos paraméter a zaj csökkentésének mértéke, a precizitási szorzó és a simítás is. Ezek együttesen határozzák meg az eredmény minőségét. Van „előzetes bemutatás”, tehát meghallgathatjuk, hogy milyen eredményre számíthatunk, sőt meg-hagyható csak a zaj is.

A végén célszerű a jelet normalizálni, a hangerő (vagy jel–zaj arány) javítása érdekében. A mentés történhet szakaszonként is, ezzel a hanganyag darabolása is megtörténhet. Az eljárással akár 60–80 dB jel–zaj arány is elérhető, ami a kiindulási 30–40 dB-hez képest nem rossz eredmény.

#### 4.2.2. Hangkazetta archiválása

##### 4.2.2.1. A magnetofon és a hangszalag

A magnetofon vékony mágnesezhető réteggel bevont műanyag szalagon rögzíti a jeleket. A jelek, mivel az átmágnesezhető réteg vékony, nagyon kis feszültséget indukálnak a letapogató rendszerben, vagyis a magnetofon fejben, mely egy speciális kemény, kopásálló bevonattal rendelkező tekercs(pár). A kicsiny indukált feszültség erősítése egy korrekciós előerősítő fokozattal kezdődik, melynek a magas hangokat kell kiemelnie. A lejátszó rendszer ezt a szalagot húzza egyenletes sebességgel a fej előtt. A szalagsebességről egy kisebb átmérőjű fém és egy rugóval rá-szorított nagyobb átmérőjű gumigörgő gondoskodik, egyes esetekben ezt a rendszert megduplázzák, a stabilabb szalagsebesség érdekében. A görgők viszonylag könnyen szennyeződnek, ami a szalag nem kívánt oldalirányú mozgásához vezet; ez károsan befolyásolja a lejátszást. A szalagsebesség egy fontos jellemzője a berendezésnek, ami 4,76 cm/s egy átlagos magnetofon esetében. Ennek beállítása referenciaszalaggal és oszcilloszkóppal történik. A gyártók egy túrésen belülré állítják ezt az értéket, tehát két azonos szalagsebességű berendezést nehéz találni. A lejátszó és felvevő kombinált fej szennyeződhet, átmágneseződhet, ami szintén minőségromlást eredményezhet. Fontos paraméter még a szalagon található sáv és a fejen található érzékelő egymáshoz viszonyított elhelyezése is. Nem megfelelő beállítással a magasabb frekvenciájú hangok mennyiségének nagymértékű csökkenése

érhető el. A gyártók ezt is határértékek közé állítják, de létezik gyárilag rosszul beállított magnetofon is.

A szalag esetén problémák a szalagra felvitt mágnesezhető réteg minőségével kezdődnek, ebből alapjában véve négyféle típust alkalmaztak. Az első a normál vasoxid, a második a króm dioxid, a harmadik a ferrokróm (ezt már nem alkalmazzák) a negyedik pedig a magnetit. A következő probléma a mágnesezhető réteg szemcsézete, egészen pontosan a mágneses információhordozók elrendezése. Ez a szalag gyártójától függően – azonos típusú mágnesezhető anyag esetén – anyagminőségben is eltérhet egymástól, tehát vannak jobb minőségű és rosszabb minőségű hangkazetták, annak függvényében, hogy a mágnesességet mennyi ideig képesek megőrizni az egyes szalagok. A harmadik probléma a műanyag réteg vastagsága, mely lejátszáskor is okozhat problémát a szalagsebesség és az oldalirányú mozgás miatt, de tárolás esetén kérdéses, hogy az egyes átmágnesezett rétegek hogyan hatnak egymásra az idő múlásával. A műanyag hosszirányban képes a megnyúlásra, ami tovább fokozza a problémákat. Továbbá a vékony szalag nagyon érzékeny a mechanikai behatásokra, mint például a gyűrődés, amely durva minőségromláshoz vezethet. Az ilyen jellegű hiányok sem pótolhatók a jel folytonosságában, még digitális módszerrel sem. Nagyon fontos szerepe van még szalagok esetén az előmágnesezésnek, ugyanis ez is potenciális zajforrás.

#### 4.2.2.2. Várható zajok

A szalagos egységek esetében az előmágnesezési zaj után a szalagnak is van egy meghatározott értékű zaja, melyet egy olyan elektronika követ, ahol több termikus zajgenerátor is működik egy időben. A zaj ezért összetett, és általában egy elég jól hallható sistergés formájában jelentkezik, tehát főleg a magasabb frekvenciákon ölt jelentős mértéket.

A magnetofon esetén alkalmazható a Dolby zajcsökkentő rendszer, de lejátszáskor csak akkor, ha a felvétel is azzal készült, ellenkező esetben nem csak a zaj, hanem a magas hangok is eltűnnek. Ha nem tudjuk egy szalag esetén, hogy használtak-e valamilyen analóg zajcsökkentőt a felvételkor, akkor a berendezés zajcsökkentőjét kapcsoljuk ki, mert így a digitális zajszűrő hatékonyabban távolíthatja el a zajt, és több hasznos jel maradhat épen.

A magnetofonnal – a gyártók ígéretei szerint – akár 60 dB jel–zaj arány is elérhető a Dolby zajszűrővel, de sajnos a szalagok minősége és a közbeiktatott erősítőfokozatok miatt a digitalizálás után ez az érték 40–45 dB-re romlik.

Mivel vannak olyan rögzítő eszközök ebben a kategóriában, amelyek a törlést és egyben az előmágnesezést egy állandó mágnessel valósítják meg, ezért az előmágnesezés szintje közel sem lesz állandó, sokkal inkább jól hallható alacsony frekvenciájú előmágnesezési zaj lesz belőle. Ez a zajforma oly mértékben elnyújtott, és véletlenszerű, hogy algoritmikus csökkentése esetenként nem is lehetséges, ugyanis a zaj mintavételezése során nem választható ki úgy zajos rész, hogy e zajból kellő mennyiségű és minőségű jusson a mintába megfelelő kivonáshoz.

#### 4.2.2.3. Zajcsökkentési eljárás

A magnetofonszalagokról érkező jelek főként magas frekvenciájú zajt tartalmaznak, mely a korrekciós előerősítő fokozatból adódik. Ez a zaj mintavételezést követően spektrális kivonással csökkenthető a zavaró mérték alá.

Csak olyan esetekben ajánlott a magnetofonszalagok digitális hanghordozóra másolása, ha az anyag bír akkora fontossággal, hogy megtegyük (más hanghordozón, nem hozzáférhető). A magnetofonszalagokról tudjuk, hogy a normál szalag sávátvitele alig több mint a fele a krómdioxidos vagy magnetit szalagokéhoz képest, és a szalag zaja, ami a mágneses információhordozó minőségéből adódik, is lényegesen nagyobb az ilyen szalagoknál.

A magnetofonnál általában nem jelentkezik impulzív jellegű zaj, ezért a zajcsökkentés elvégezhető egyetlen lépésben. A digitalizált jelből (az elejéről vagy a végéről) zajmintát veszünk, majd pedig spektrális kivonással távolítjuk el a zajt. Az eredmény – bár az eljárás egy lépésből áll csupán – nem lesz olyan minőségű, mint a hanglemezek esetében, ugyanis a magasabb frekvenciájú hangokat a rögzítő rendszer és a szalag már korábban eltüntette, ezért lejátszáskor sem voltak már jelen. Ezek a hangok már semmilyen eszközzel nem állíthatók vissza.

## 5. Összegzés

Dolgozatom célja volt a környezetünkben előforduló zajok osztályozása, és azok csökkentésére megoldást találni. A potenciális zajforrások nagytöbbségére a dolgozatom első szakaszában világítottam rá. Tekintettel arra, hogy zaj hasonlóan az információt hordozó jelhez véletlenszerű, ezért modellezése és csökkentése közel sem olyan egyszerű, mint ahogyan azt a köznapi ember gondolja. A zajok modellezése a zajcsökkentés szempontjából jelentős, ugyanis egy megfelelő modell segítségével az eljárás könnyebben algoritmizálható, és alkalmazható a digitális jelfeldolgozásban. A számítógépek miniatürizálásával és gyors fejlődésével korunkban egyre nagyobb teret kap a kommunikációs és jelrögzítő rendszerekben a digitális jelek alkalmazása, mely a feldolgozhatóság és a hozzáadódó zajokkal szembeni jó „ellenálló-képességének” köszönheti sikerét.

A zajok modellezését kétféle szempont szerint közelítettem meg dolgozatomban, melyekből kiderült, hogy a véletlenszerű jelek előre meg nem jósolható értékeket is felvehetnek. A modellek alkalmazhatók más köznapi véletlenszerű események modellezésére is, melyek messze esnek az informatika tudományterületétől, mint például a közgazdaságtan, vagy a meteorológia, vagy az orvostudomány.

A probléma megoldása a zajcsökkentési eljárások alkalmazása. Többféle módszer létezik, melyekből csak néhányat volt alkalmam kiemelni, dolgozatom szűkös keretei miatt. A módszereket az adott problémának megfelelően, jelen esetben a zajok fajtáihoz illően kell alkalmazni, a várt eredmény elérése érdekében. A zajok csoportosítása, és forrásuk meghatározása ezért nagyon fontos, a legalkalmasabb eljárás kiválasztása szempontjából. Egy véletlenszerű jelsorozatból bizonyos esetben nehéz eldönteni még a szétválogatás után is, hogy melyik a számunkra hasznos információt hordozó jel. Erre nagyon szemléletes példa egy gépjármű motorja, mely az utas szempontjából zajforrás, de egy hozzáértő szerelő sok értékes információhoz juthat a hallott hangokból.

Az utolsó szakaszban az elméleti ismeretek gyakorlatba ültetését próbáltam kifejteni, saját tapasztalataim alapján. Több mint tíz évvel ezelőtt volt alkalmam több „házi stúdióban” készült könnyűzenei demófelvételhez segítséget nyújtani, és jómagam is aktív részese voltam egyes felvételek elkészítésének, mint hangszeres „zenész”. Jelenlegi iskolámban néha alkalmam nyílik kamatoztatni korábbi ismer-

reteimet, és dolgozatom megírásával nagyobb rálátást nyerhettem a későbbi munkámhoz a zajok forrásaira és megfelelő mértékű csökkentésükre.

A régebbi hanghordozókon tárolt, többnyire zenei felvételek korszerű adathordozóra mentése több ismerősömnek eszébe jutott már, de megfelelő eszköz hiányában maguk nem tudták elvégezni a műveletet. Érdeklődési körömbé tartozik a hangrendszerek felépítésének, és működésének megismerése és tanulmányozása, ezért több, mikrobarázdás lemez lejátszására alkalmas eszköz és magnetofon is található otthonunkban, mellyel a hanganyagok digitalizálásához forrást tudok biztosítani. A számítógépben található hangkártyákra mindig figyeltem, hogy ne csak gyenge minőségű hangjelet tudjon produkálni a komolyabb hangrendszer számára, hanem megfelelő működésével minőségi zenehallgatásra legyen alkalmas.

A dolgozat megírása során törekedtem arra, hogy a címben megragadott témát minél jobban, és több szempont szerint is kifejthessem, amit végeredményében sikerült megvalósítanom.

## 6. Irodalomjegyzék

1. Saaed V. Vaseghi – Advanced Digital Signal Processing and Noise Reduction  
Second Edition  
Copyright © 2000 John Wiley & Sons Ltd.
2. Valószínűségszámítás és statisztika előadások anyaga  
Nagy Márta – 2001–2002 tanév
3. Molnár-Sáska Gábor – Rejtett Markov modellek statisztikai vizsgálata  
Tézisfüzet – 2005  
forrás: [http://doktori.math.bme.hu/Molnar-Saska\\_tezisfuzet.pdf](http://doktori.math.bme.hu/Molnar-Saska_tezisfuzet.pdf)
4. dr. Tolnai János – Rádióamatőr vizsgára felkészítő tananyag  
Puskás Tivadar Távközlési Technikum, 2006.  
forrás: [http://www.puskas.hu/r\\_tanfolyam/r\\_tananyag.html](http://www.puskas.hu/r_tanfolyam/r_tananyag.html)