



Budapesti Műszaki és Gazdaságtudományi Egyetem

Hálózati Rendszerek és Szolgáltatások Tanszék,

Akusztikai és Stúdiótechnikai Laboratórium (LAST)

---

Dr. Márki Ferenc

# Digitális hangfeldolgozás

Budapest, 2002. január

átdolgozva: 2007. április  
2013. április  
2014. október

# Tartalomjegyzék

|                                                                         |                              |
|-------------------------------------------------------------------------|------------------------------|
| <b>TARTALOMJEGYZÉK</b>                                                  | <b>1</b>                     |
| <b>ELŐSZÓ</b>                                                           | <b>3</b>                     |
| <b>1. HANGRÖGZÍTÉS</b>                                                  | <b>5</b>                     |
| RENDSZERÖSSZEÁLLÍTÁS                                                    | ERROR! BOOKMARK NOT DEFINED. |
| <b>2. DIGITÁLISAN TÁROLT HANGANYAG SZERKESZTÉSE</b>                     | <b>8</b>                     |
| <b>2.1. ELEKTROMOS JELLEGŰ ÉS FELVÉTELTECHNIKAI HIBÁK KIKÜSZÖBÖLÉSE</b> | <b>8</b>                     |
| (KI-)VÁGÁS                                                              | 9                            |
| MEGJEGYZÉS MANIPULÁLANDÓ/TÖRLENDŐ SZAKASZOK KIJELELÉSÉHEZ               | 10                           |
| DC OFFSZET                                                              | 10                           |
| HANGERŐ SZABÁLYOZÁS                                                     | 13                           |
| FADE-IN, FADE-OUT                                                       | 14                           |
| KÉZI BEAVATKOZÁS - HULLÁMFORMARAJZOLÁS                                  | 15                           |
| <b>2.2. VÁGÁSI MUNKÁLATOK</b>                                           | <b>17</b>                    |
| CROSSFADE                                                               | 19                           |
| TEMPÓ VÁLTOZTATÁS                                                       | 19                           |
| HANGMAGASSÁG VÁLTOZTATÁS                                                | 20                           |
| <b>2.3. EFFEKTEK</b>                                                    | <b>23</b>                    |
| HANGSZÍNSZABÁLYOZÁS                                                     | 23                           |
| ZAJZÁR                                                                  | 24                           |
| DINAMIKA-KOMPRESSZIÓ                                                    | 27                           |
| LIMITER                                                                 | 31                           |
| VISSZHANG, ZENGETÉS                                                     | 31                           |
| KÓRUS                                                                   | 35                           |
| LEBEGTETÉS (FLANGER)                                                    | 36                           |
| FÁZISTOLÓ (PHASER)                                                      | 38                           |
| VIBRÁLÁS                                                                | 39                           |
| TREMOLÓ                                                                 | 39                           |
| SZTEREÓ HATÁS                                                           | 39                           |
| TORZÍTÓ                                                                 | 40                           |
| <b>3. KEVERÉS</b>                                                       | <b>42</b>                    |
| <b>3.1. ELŐKÉSZÜLETEK</b>                                               | <b>43</b>                    |
| <b>3.2. LEJÁTSZÁS INDUL – PC-VEL!</b>                                   | <b>45</b>                    |
| <b>3.3. KEVERÉS</b>                                                     | <b>46</b>                    |
| ANALÓG KEVERŐ JELÚTJAI                                                  | 47                           |
| VIRTUÁLIS KEVERŐK                                                       | 51                           |
| AUTOMATIKUS VEZÉRLÉS                                                    | 58                           |
| <b>4. HANGTÖMÖRÍTÉSI ELJÁRÁSOK</b>                                      | <b>59</b>                    |

|                                                                                    |           |
|------------------------------------------------------------------------------------|-----------|
| <b>4.1. SPEKTRÁLIS ELFEDÉSI JELENSÉGEK []</b>                                      | <b>61</b> |
| <b>4.2. IDŐBELI ELFEDÉSI JELENSÉGEK []</b>                                         | <b>65</b> |
| <b>4.3. BITKIOSZTÁS [1, ]</b>                                                      | <b>67</b> |
| <b>4.4. MPEG AUDIO</b>                                                             | <b>72</b> |
| MPEG-1 AUDIO [5]                                                                   | 72        |
| MPEG-2 AUDIO [6]                                                                   | 76        |
| <b>4.5. AC-3</b>                                                                   | <b>80</b> |
| <b>4.6. AZ AC-3 ÉS AZ MPEG-2 TÖBBCSATORNÁS KÓDOLÁSI ELJÁRÁSOK ÖSSZEHASONLÍTÁSA</b> | <b>86</b> |
| <b><u>IRODALOMJEGYZÉK</u></b>                                                      | <b>89</b> |

## Előszó

E jegyzet anyaga elsősorban zenét kedvelő, mérnöki ismeretekkel rendelkező hangmérnökjelölteknek íródott, akik a műszerek és stúdiótechnikai berendezések elektronikai ismeretein túl a zenekészítés szempontjából lényeges gyakorlati ismeretekre is szert kívánnak tenni. Bár Magyarországon ma, 2002-ben hangmérnökképzés, mint felsőfokú oktatási szak, nem létezik, az igény mégis megvan rá. Végzős mérnökök és zenekedvelők egyaránt vágnak bele ebbe a szakmába több-kevesebb sikerrel. E jegyzet az imént említett két fajta ember gondolkodásmódját próbálja egyesíteni. Fontos hangsúlyozni, hogy a „jó hangmérnök” mind a két témában járatos. A mérnöki ismeretek mindenképpen szükségesek ahhoz, hogy a berendezések működését megérthessük, átláthassuk. Ez azért kiemelt fontosságú, mert a számítástechnika rohamos fejlődésével a hangtechnikában is egyre sokoldalúbb – a digitális technikára jelentős mértékben támaszkodó – eszközök jelennek meg a piacon, melyeket csak akkor tudunk megfelelően használni, ha általános, jól megalapozott szemléletmóddal rendelkezünk, és műszaki alapismereteinkre támaszkodhatunk. A zenei ismeretek szempontjából itt nem is zenei műveltségre, hanem inkább zenei érzékre van szükség. Annak se reménytelen a pályája, aki nem tanult hosszú éveken át valamilyen hangszeren, avagy járt szolfézásra, bár sokszor könnyebb a helyzet, ha legalább a hangszerekről alapismeretekkel rendelkezünk. A felvételkészítés szempontjából azonban a legeslegfontosabb, hogy rendelkezünk megfelelő hallással, képezzük a hallásunkat: járjunk koncertekre, hallgassunk felvételeket és főként gondolkozzunk el rajta, hogy mitől jó egy-egy előadás, és végül ami elengedhetetlen, az a zenei képzelőerő. E nélkül ui. ötletszegények, laposak lesznek a felvételeink, amelyek csak mások által kidolgozott hangzásképeket utánoznak le, avagy ollóznak össze.

A jegyzet készítése során igyekeztem a felvétel-készítés lépéseinek sorrendjében felépíteni a tematikát. Azt hiszem óhatatlanul is érződik, hogy könnyűzenei felvételek készítésében vagyok jártas, de hála Dr. Kovács Lóránt barátomnak, a komolyzenei felvételek készítésének szempontjairól és kivitelezéséről is tudtam helyenként említést tenni. Igyekeztem a

hangtechnikában szokásos nyelvezetet használni és törekedni az általában angolul használt kifejezések bevett magyar megfelelőit alkalmazni. Említést érdemel a zene/szöveg/hanganyag/darab/szám szóegyüttes, melyek általában mind ugyanazt jelentik: a felvenni, feldolgozni kívánt zenei anyag. A jegyzet jelentős része szoftveres megvalósítások használatáról szól, amelyek tárgyalásmódjában óhatatlanul befolyásoltak az általam használt szoftverek. Ezek megnevezésére nem szeretnék külön kitérni, hiszen e jegyzet semmilyen reklám vagy propaganda célt nem kíván szolgálni. Itt is az eszközök képességeire szeretném felhívni a figyelmet, nem pedig a konkrét megvalósítások kezelőfelületeire vagy apró részleteire. Ezért elnézést kérek azoktól az Olvasóktól, akik egyes leírásokat túl rövidnek találják és bizonyos gombok, kapcsolók leírását nem lelik. Remélem azonban, hogy az általános alapokat felhasználva mindenki képes lesz a különböző szoftveres és hardveres megvalósítások kezelését hamar elsajátítani és nem okoz gondot az egyedi kezelőfelületek újdonsága.

A jegyzet kidolgozásában nagy mértékben támaszkodtam a Budapesti Műszaki Egyetemen szerzett mérnöki ismereteimre, melyekkel 4 éve vágtam bele a házi, számítógépes hangfelvétel-készítésnek. Külön köszönetet szeretnék mondani zenésztársaimnak, a Nut Rockers együttesnek, akikkel 10 éve zenélünk közösen. Ők segítettek hozzá a felvételkészítés gyakorlati tapasztalataihoz, köszönet a türelmükért és a kezdetben nehézkes szoftverkezelés okozta kényelmetlenségek elviseléséért, melyekkel az első CD-nk elkészüléséhez hozzájárultak.

Budapest, 2002. január

Márki Ferenc

## 1. Hangrögzítés

Ez a fejezet igyekszik gyakorlati tanácsokat adni hangfelvétel elkészítéséhez, legyen az klasszikus avagy könnyűzenei felvétel, stúdió, házi avagy helyszíni rögzítéssel. A teljes hangprodukció készítés lépései nagyon általánosan:

1. Rendszer-összeállítás, jelszintek beszabályozása
2. Felvétel
3. Szerkesztés: vágás és utómunkálatok
4. Keverés és effektezés
5. Kész hanganyag tárolása

Ezek közül egyik másik lépés olykor elmaradhat, mégis a kép akkor teljes, ha mindegyikről esett szó.

### Rendszer összeállítás

Különös gondot kell fordítanunk a rendszer felépítésére, hiszen a felvétel minőségét döntően befolyásolja, hogy milyen rendszeren „küzdötte” át magát a „zene”, mire digitális formában rögzítésre került. Először is tisztában kell lenni azzal, hogy a különböző eszközök kimenetei különböznek-e valamiben és ha igen, akkor miben. A legelső kérdés, hogy analóg avagy digitális kimenetről van szó. Ha analóg, akkor bár az eszközök nagyon sok félek, a kiadott jelek szerencsére két nagy csoportba sorolhatók:

- mikrofon szintű és
- vonali szintű

Egyértelmű, hogy a mikrofon szintű jelek közé a mikrofonok tartoznak, az összes többi pedig elméletileg a vonali szintűek közé tartozna. Vannak azonban gyengébb jelforrások, amelyek bár erősebbek, mint a mikrofonok, mégis a szintjük sokszor nem közelíti meg a minimum 200-300 mV-os szintet, amelyet

már minden keverő vonali szintnek képes kezelni. Ilyen források lehetnek pl. a gitárok, a szintetizátorok. Érdekes azonban ezekkel is első körben a vonali bemeneten próbálkozni, mert sokszor az is tud eleget erősíteni a jelen. A mikrofon bemenetek pedig általában lényegesen többleterősítéssel rendelkeznek, ami többlet zajt hozhat a rendszerbe, és ha nem figyelünk eléggé, akkor torzítást is. Előfordulhat ui., hogy vonali szintű jelet kötünk a mikrofon bemenetre (mert pl. „a bekötendő kábel csatlakozója a mikrofon bemenetbe illett bele”), és anélkül, hogy túl hangosnak találnánk – mert a csatornahangerő-szabályzót kellően lehúztuk – mégis úgy érezzük, hogy valami nem stimmel. A gond ott lehet, hogy pl. kismértékben túlvezéreljük az előerősítő fokozatot és erre esetleg nem figyelünk fel egyből. Természetesen ha van kivezérlésjelző a szintezés után, akkor könnyű a beállítás, de sokszor kell olyan eszközzel dolgozni, amin nincs bepillantási lehetőség az előerősítő utáni jelbe. Igyekezzünk mindig nagy szintű jeleket továbbítani, tehát ha egy eszköz kimenetén a hangerő változtatható, akkor egy jó erős jelet állítsunk be. Azért nem mondok a „lehető legnagyobb”, mert sajnos sok olyan eszköz van a piacon, ami maximális hangerő mellett már torzít. A megfelelő szint megválasztásánál az is fontos szempont, hogy az adott hangforrás vajon milyen mértékben változtatja majd meg a hangerejét. Mivel a felvételt digitálisan fogjuk rögzíteni, ahol is a legdurvább hibát az analóg-digitál átalakítók túlvezérlésével követhetjük el, ezért készüljünk fel nagyobb beütésekre, kiabálásokra, és egyéb impulzus szerű hangerő löketekre.

Analóg eszközöknél különös gondot fordítsunk a felhasznált kábelek és csatlakozók minőségére; ezen a területen többszörösen megtérül a minőségi eszközökre való többletráfordítás – nincs annál rosszabb, mint egy „kábelrengetegben” keresni a kontakthibát, avagy azt nyomozni, hogy melyik kábel szedi fel a „brummot”.

Digitális eszközöknél kicsit más a helyzet, bár ott nem kritikus a csúcsmínőségű kábel használata, hiszen a digitális technika egyik legfontosabb előnye pl. a hibamentes kódrekonstrukció: tehát ha az adó által kiadott jel egy kevés zavart felvesz, attól még a vevő általában képes a hibamentes dekódolásra. A végtelenségig azonban nem! A digitális rendszer kiépítésénél leginkább arra kell figyelmet fordítanunk, hogy a teljes rendszer csak egyetlen órajelet

használjon. Vagyis egy eszköz dolgozik master üzemmódban (ő az, aki az órajelet generálja és adja tovább), az összes többi pedig slave üzemmódban kell, hogy dolgozzon: nem a saját órajelet, hanem a kívülről érkező órajelre szinkronizál. Ez érkezhethet az adatjellel együtt (egy kábelen, megfelelő kódolással), de érkezhethet külön kábelen is.

Ha a rendszert fizikailag összeállítottuk és analóg jeleknél a szinteket beállítottuk, akkor indulhat a felvétel. Bár ez sokszor unalmas lehet a hangmérnök számára, ha az előadó nem túl jó, de viszont jót akar felvenni, mégis a felvétel sokszori meghallgatása során „meg lehet jegyezni” a zenét és az utómunkálatok során kész elképzelésekkel dolgozni.



## 2. Digitálisan tárolt hanganyag szerkesztése

A leírás igyekszik általános szinten maradni, a feldolgozási módszereket a zenei hatásokon túl mérnöki megközelítéssel is ábrázolni. Egyes részletek olykor lényegtelennek tűnhetnek, de olyan eszköztárat kívánok az olvasó elé tárni, mely a hangtechnikailag gyengén felszerelt hangfelvételt készítőik számára is használható. Bizonyos eszközök, módszerek egy profi hangstúdióba bejáratos ember számára feleslegesek, de érzésem szerint a többség nem így kezdi. Bár az itt bemutatandó lépések céleszközökkel valósítódnak meg a professzionális hangtechnika világában, a számítógépes szemlélet azonban – némileg alacsonyabb szinten ugyan, de – itt is használható. Induljunk ki tehát abból, hogy a zenei produkció megtörtént, a hanganyag rögzítésre került, és minden szem, ill. fül a „hangmérnökökre” terelődik. Innentől az ő munkája határozza meg, hogy a jóból kihozza a maximumot, avagy a nagyon jót elrontja.

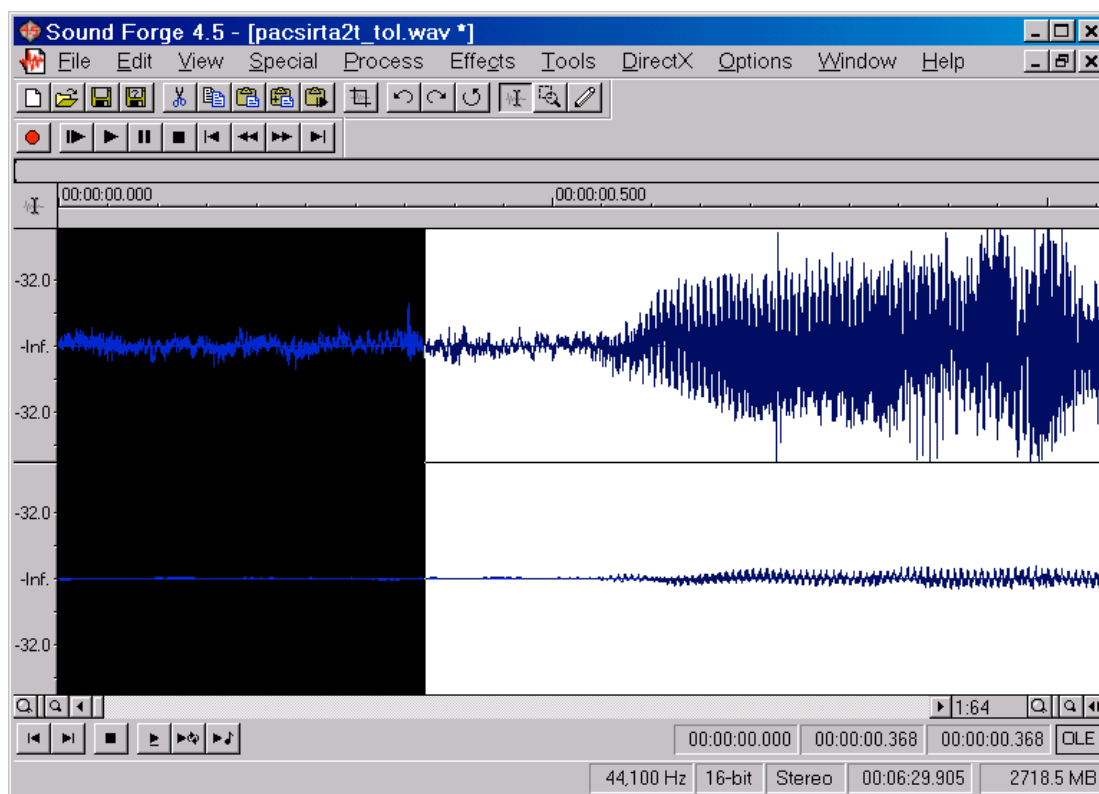
Klasszikus zenét hallgatók körében sokszor negatív érzéseket vált ki a „zenei anyag számítógépes feldolgozása, szerkesztése” és közben nem is sejtik, mennyi utómunkálaton ment keresztül kedvenc klasszikus CD-jük is. Feldolgozás alatt nem kell feltétlenül manipulálást, vagyis az eredeti előadás „meghamisítását” érteni. Bár a könnyűzenében teljesen megszokottá vált, hogy alakítjuk, színezzük a zenét, mégis létezik sokkal alapvetőbb feldolgozás is.

### 2.1. Elektromos jellegű és felvételtechnikai hibák kiküszöbölése

Egy mérnök számára kifogástalan berendezés nem létezik, csak olyan, ami az adott specifikációban megjelölt hibakorlátokon belül marad. Így van ez a hangtechnikában is, minden eszköznek van hibája: linearitási hibák, alapzaj, frekvenciamenet egyenetlenségek, stb. Ekkor még nem beszéltünk a gyengébb minőségű hangszerekről, a tapasztalatlan felvételre éneklő előadóról és a rendszerbe be-becsúszó technikai malőrökről. Ezen hibák kijavítása a felvételt ugyan meghamisítja, de a produkciót nem. Ezen hibák következtében ugyanis a hanghordozón tárolt anyag sohasem egyezik a valóságban elhangzottal – csak maximum elég jól követi. Ezért – különösen a technikai jellegű – hibák kijavítása csak közelebb viszi a felvételt a valósághoz.

## (Ki-)Vágás

A felvételeket szeretjük biztonsággal a hangesemény megkezdése előtt indítani, és kellő ráhagyással leállítani, hogy biztosan ne maradjon le semmi a kicsengésből sem. Ezzel szemben lehallgatáskor – különösen a CD lemezek világában – nem szeretjük, ha hosszas szünetekkel kezdődnek a számok (1-2 másodperc) ill. a kicsengés és az utána következő csend nagyon elnyúlik. Ezért a további munkálatok során feldolgozandó állományt rögtön az elején célszerű megszabadítani a felesleges eleje-vége szakaszoktól. Ehhez először meg kell keresni a hanganyag elejét és végét, célszerűen 10-100 ms ráhagyással. Minden esetben ellenőrizzük meghallgatással is, nem marad-e le valami, amire szükség volna. Ezután kijelöljük a törlendő szakaszt (ld. a 2.1. ábrát) és a törlés paranccsal eltávolítjuk. (Ne csodálkozzunk, hogy hosszú hanganyag elejének megvágása sokáig tart, hiszen ilyenkor a teljes fájl elejét el kell hagyni, és az összes adatot előrébb kell másolni. Legalábbis a jelenlegi Microsoft operációs rendszereknél.)



2.1. ábra. Kijelölés

## Megjegyzés manipulálандó/törlendő szakaszok kijelöléséhez

Amennyiben nem a teljes hanganyagon dolgozunk, hanem csak egy kisebb szakaszát jelöljük ki, úgy számolni kell azzal, hogy ha környező hanganyaggal való illeszkedési pontokon „törést” okozunk akkor az hallható pattanást okoz a lejátszás során (bővebben ld. a következő fejezetet). A törés fogalma nehezen definiálható, de a gyakorlott szem könnyen észreveszi: a jelgörbe meredeksége hirtelen változik. Ezt a legkönnyebben úgy lehet elkerülni, hogy a kijelölések elejét és végét 0-átmenetekhez tesszük, vagyis ahol a jel átmegy a 0 szintegyenesen. Amennyiben a kijelölt darabkát eltávolítani szándékozunk, úgy érdemes arra is figyelni, hogy a kivágott szakaszt megelőző és az utána következő jelek egyike alulról tartson a 0 felé és ne lépje túl, a másik oldal pedig felülről. A hanganyag elejének és végének törlésekor azt kell figyelembe venni, hogy alapértelmezésben a számot teljes csenddel záródó szám előzi meg és teljes csendről induló szám követi.

Hangerőmódosulással járó beavatkozás (hangerő- / hangszínszabályozás, dinamikamódosítás, stb.) esetén a kijelölt szakasz határán eredetileg folytonos jel sérülhet, ha a kijelölt a tartomány első hangmintája túl nagy mértékben megy el a helyéről. Ha a nagy mértékű beavatkozás elkerülhetetlen, akkor keressünk halk szakaszt, esetleg az átmenet 1-2 hangmintáját kézzel módosítsuk.

Sztereó (többcsatornás) anyagnál gondot okozhat, hogy a két csatorna többnyire sosem megy át egyszerre a 0-án. Ilyenkor sajnos kénytelenek vagyunk a mérnöki intuíciónkra hagyatkozni és a „lokálisan optimális megoldást” érzésre meghatározni.

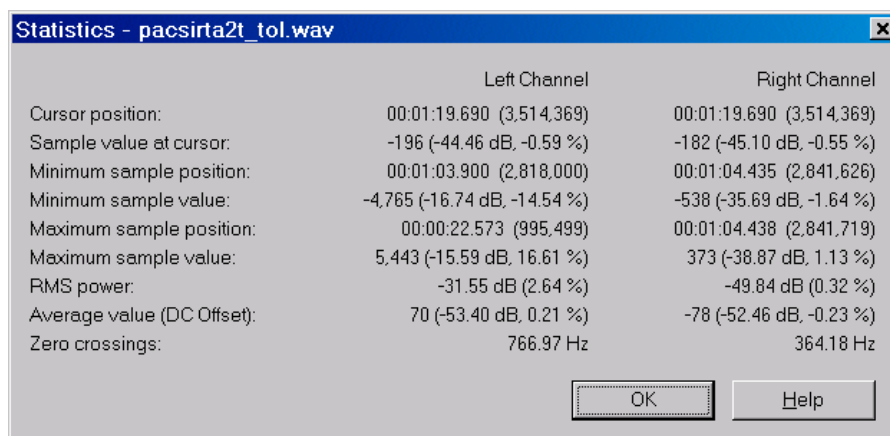
## DC offset

Analóg eszközök aligha küszködnek ezzel a problémával, a csatolókondenzátorok ui. mindenhol – szinte – DC mentessé simítják a jelet. Ezzel szemben az analóg-digitál átalakítók többnyire DC csatoltak, tehát a közvetlen megelőző fokozat DC hibáját rögzítik. Gond akkor jelentkezik, amikor a DC szintben ugrás történik. Pl. azért, mert a lejátszás során egymást követő számok nem azonos erősítési paraméterek mellett digitalizálódtak. Még rosszabb, ha nagyobb jelugrás hatására a DC szint „úszik”, pl. ha az átalakítót megelőző

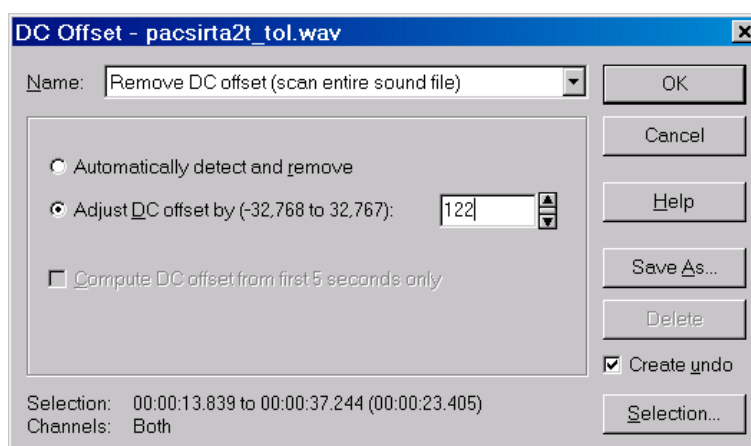
fokozat nagyon kis frekvenciákig lemegy (0.1 Hz nagyságrend), ilyenkor ui. a jel egy nagyon hosszú periódusidejű jelre (10 másodperc) van ráültetve és ennek hatására a DC szint úszni látszik. Akkor is adódhat DC ugrás, ha az egyhuzamban felvett anyag bizonyos részeit hangosítom – pl. a zenész felkonferálása – és így a jellel együtt a DC szintet is változtatom. A DC ugrás – levezethető módon – nagyfrekvenciás komponensek megjelenését okozza, azaz pattanás hallható. A mértéke természetesen a DC ugrás mértékétől függ, hogy meghallható-e, az pedig a fülünktől és a pattanást követő hangeseménytől egyaránt. Ha ui. nagyon rövid időn belül (1-2 ms) elegendően hangos jel következik, akkor előelfedés lép fel: a hangos jelet megelőzően rövid időre megemelkedik a hallásküszöbünk. Erről a jelenségről később a „Hangtömörítési eljárások” című fejezetben olvashatunk.

A DC offset eltávolítása rendkívül egyszerű feladat, a teljes DC mentesítendő szakasz hangmintáinak matematikai átlagolásával megkapjuk a DC szintet, majd ezt az értéket kivonjuk minden egyes hangmintából. Egy dologra kell csak figyelni, hogy a legnagyobb amplitúdójú hangminta növelése túlcsoordulást okozhat. Ez pedig már az analóg-digitál átalakítás rejtelseibe bevezetés során kifejtett okok miatt kerülendő. Ilyenkor vagy halkítsuk le először a felvételt, majd azután számoljunk DC offset-et és alkalmazzunk korrekciót, vagy törődjünk bele, hogy nem távolítjuk el, csak csökkentjük a DC mértékét, azaz a megengedhető maximális értéket alkalmazzuk a korrekció során.

Amint arról szó esett, manuális DC eltávolítás során szükségünk van a hanganyag avagy annak egy darabjának egyes információira. Ezt többnyire statisztika címszó alatt fut a különböző hangszerkesztő programokban. Minket jelenleg a maximális és a minimalis hangmintaértékek érdekelnek, illetve a DC offset. A megszokott decibeles gondolkodásmód helyett most a konkrét lineáris számértékeket kell használni. A 2.2. ábrán látható példa esetében a bal csatornán pl. 70-et kell minden értékből levonni, gond tehát a minimális értékű hangmintá(k)nál lehet. Ez jelen esetben -4765, azaz a -32768-at meg sem közelítő érték. A jobb csatornán 78-at kell hozzáadni a hangmintákhoz, a maximális hangminta értékét nézve (373) pedig itt sincs gond, a manuális DC mentesítés mindkét csatornán végrehajtható.



2.2. ábra. Statisztika párbeszédablak



2.3. ábra. DC offset eltávolítása párbeszédablak

A 2.3. ábrán egy tipikus DC eltávolító vezérlőpanelt látunk. A manuális üzemmód esetében csak a megfelelő DC offset-et kell beírunk. Automatikus üzemmódban még ennyi dolgunk sincs, ilyenkor viszont – előzetes számítás megtakarítás céljából – megadható, hogy a kijelölt (avagy teljes) hanganyagnak csak egy rövidebb darabkája (jelen esetben az első 5 másodperc) alapján számoljon DC szintet a program. Ez lassabb processzorok illetve stabil (nem úszó) offset-tel rendelkező berendezések használata esetén alkalmazható opció.

Megjegyzés 2014: mára a hangkártyák DC offset-je egyre inkább elhanyagolható, ezért a szoftverek egyre kevésbé támogatják a DC offset kézi manipulálását. Ne csodálkozzunk, ha esetleg egyáltalán nem találkozunk a paranccsal avagy csak nagyon elrejtve valahol. Sokszor pl. a „Normalize” funkcióval kötik egybe, ahol egyszerűen csak beállíthatjuk, hogy milyen DC szintre szeretnénk hozni a hanganyagot. Itt nyilván csak a 0.0 % jön szóba...

## Hangerő szabályozás

Talán nem is kellene külön szót említeni róla, annyira természetes fogalom. Mégis mérnöki megközelítéssel egy-két dolgot érdemes hozzáfűzni. Triviális, hogy részletek kiemelésére, esetleg visszaszorítására ill. hosszútávú szintkülönbségek kiküszöbölésére használandó. Az első két esetben mindenképpen a fülünkre hagyatkozva a harmadiknál azonban hasznos támpontot nyújthat a statisztika párbeszédablak. Az itt található RMS érték a hangteljesítményről ad információt. Bár ez az adat nem áll egyértelmű kapcsolatban a szubjektív hangosságérzettel, melyet pl. a hangszín jelentősen befolyásol, mégis irányjelzőként, különösen azonos típusú zenéknél jól használható. A körülbelül azonos RMS szintre hozott hanganyagok nem fognak túlzott hangerőváltozás érzetet okozni. Hangerő szabályozás esetében mindenképpen figyelni kell, nehogy túlcsoportulást okozzunk – kritikus esetekben ellenőrizzük a statisztika paranccsal –, de érdekes tudnivaló az is, hogy a pontosan 2x, 4x, stb. (dB-ben 6.02, 12.04, stb.) erősítés nem hoz további kvantálási zajt a rendszerbe, hiszen pontosan 1, 2 stb. bittel toljuk el a hangminták értékét. Kellően nagy jelszintek esetén persze nem gond a kismértékű kvantálási zaj, de halk részleteknél már kitűnhet. A halk alatt itt a számábrázolási tartomány aljához közel álló értékekre kell gondolni.

*Példa:* Ha 16 bites a számábrázolás, akkor 32767 a pozitív maximális érték, 0 pedig a teljes csend. Ez a számábrázolás 90.3 dB-es dinamikatartományt jelent, melyből kb. 10-20 dB túlvezérlési tartalék gyanánt lejön. (Úgy állítjuk be a felvételt, hogy a hangos jeleneteknél is legyen 10-20 dB még a teljes kivezérlésig.) Zenéről lévén szó, a hangteljesítmény és a csúcserő kb. 4-5x-ös arányban állnak, vagyis a hangteljesítmény a csúcs alatt kb. 12 dB-el helyezkedik el. Így nagyjából 65 dB-re adódik az átlagosan hangos részlet energiája. Ennél mondjuk 35 dB-el halkabb részletre számítva a piano úgy 30 dB-es teljesítményű, azaz 32 környéki (hangminta-) értékeknél járunk. Ehhez képest az átlagosan max. fél bitnyi hiba – pontos számítás mellőzésével – úgy 30-40 dB-es jel/zaj viszonyt eredményez – ez pedig nem túl sok.

## Fade-in, fade-out

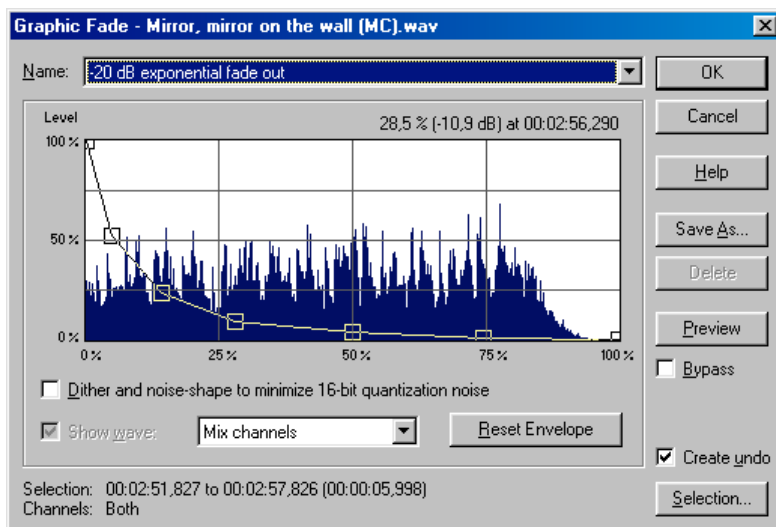
Felhangosodásnak és elhalkulásnak lehetne leginkább fordítani ezeket a kifejezéseket. Leggyakoribb felhasználási területük a számok eleje illetve vége. Mivel (majdnem) minden felvétel zajszintje megjelenik az ábrázolható jeltartományban, ezért a számok elején a semmiből hirtelen ugrana fel a zajszint. Hogy ez egy kicsit természetesebben hangozzék, célszerű felhangosodást alkalmazni. A Fade-in időtartalma relatív rövid lehet, többnyire 5-10 ms elegendő. A szám korrekt lezárása érdekében szintén 0-ra, a teljes csendre kell megérkezni, ezért ugyanaz az eljárás, mint a szám elején, a végén is használható csak épp a fordított irányba. A neve Fade-out. Ennek hossza a záróakkord lecsengése alapján határozandó meg, nagyságrendileg 100-200 ms.

Másik felhasználási területük az, amikor kifejezetten hosszú időtartamon (akár másodperceken át) használunk elhalkulást vagy hangosodást. Pl. egy koncerten ünneplő tömeg hangulatának visszaadásánál szeretjük a hosszú „beúsztatásokat”, avagy a nagyon lassú elhalkulást.

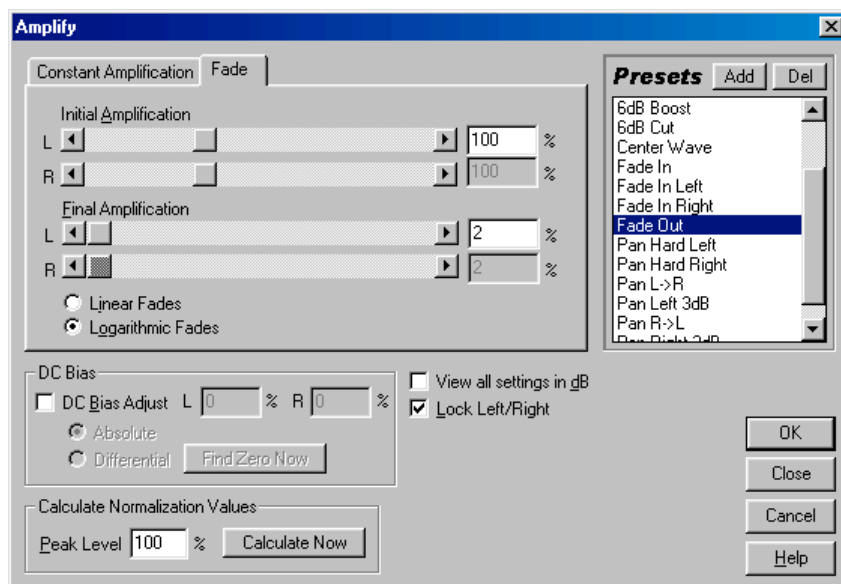
Előreteként az „effektezés” fejezetbe megemlítendő, hogy a halkulást akkor is lehet használni, amikor egy akkord kicsengése a sávonként rögzített hangszereken jelentős eltérést mutat. Pl. a basszus gitár alig halkul míg az összes többi hangszer szinte teljesen kicseng. Ilyenkor kicsit „rásegíthetünk” az elhalkulásra. Megfelelő lecsengési görbe beállításával nagyon kellemes és természetes hatást tudunk elérni.

A fade-in, fade-out többnyire a legegyszerűbb hangszerkesztő programokba is beépülésre kerülnek. Ilyenkor általában olyan a megvalósítás, hogy egy a 0 %-ot a 100 %-al összekötő egyenest rajzolnak a kijelölt szakasz fölé és ezzel a hangerőváltozással szorozzák végig a hangmintákat – egyenként. Komolyabb programok esetében különféle opciókat találunk az alapértelmezett lineáris fade-in/out mellett. Sokszor megadható, hogy ne a 0 %-ra, hanem valami egyéb szintre történjen a halkulás. Ezzel mintegy lehetőséget kínálva a lecsengési görbék megfelelő meredekségének kényelmes beállítására. Ennél viszont sokkal fontosabb képesség az, hogy a görbe alakja megválasztható legyen. A lineáris lecsengés csak matematikailag jó (és számítástechnikailag könnyű) megoldás, a

fülünk szempontjából az exponenciális lecsengés, azaz a *decibelben* lineáris lecsengés a jó. Ennek a megvalósítása azonban számításigényesebb feladat, ezért sokszor lineáris szakaszokkal közelített exponenciális görbét alkalmaznak. Ez nem okoz gondot rövid lecsengések esetén, hosszú lecsengéseknél azonban túlságosan elnyúlnak a lineáris szakaszok és újra a nem természetes hangzásnál vagyunk. Sokszor arra is lehetőségünk nyílik, hogy a lecsengési görbét magunk rajzoljuk meg; a megadott pontok közötti többnyire lineáris átmenettel.



2.4. ábra. Lineáris szakaszokból megadható tetszőleges alakú lecsengés



2.5. ábra. Kezdeti- és végerősítés megadása, lineáris és exponenciális karakterisztika

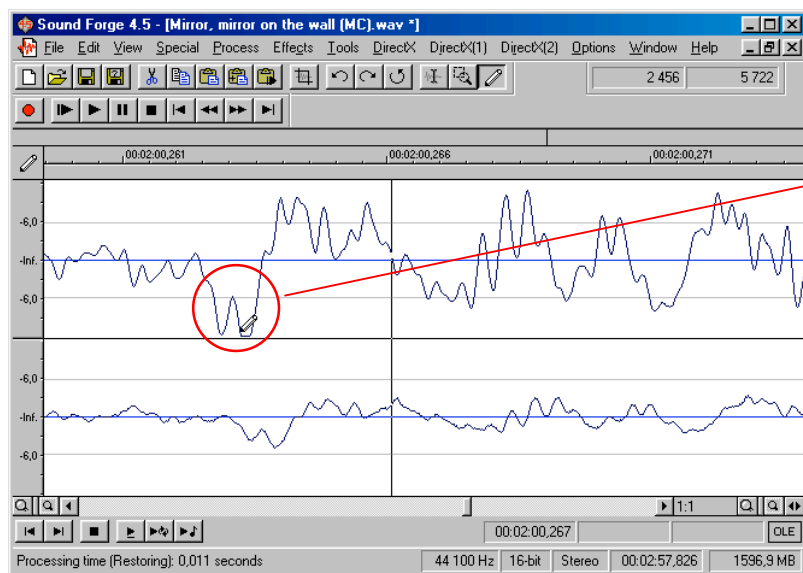
## Kézi beavatkozás – hullámforma-rajzolás

Bizonyos esetekben szükség lehet a hangminták közvetlen – kézzel való – módosítására. Ez többnyire akkor használatos, ha a hiba kis mértékű, és



egyedinek tekinthető. Ilyenkor nem érdemes zaj/zavarszűrő algoritmust konfigurálni a pontos eredmény elérésének érdekében, ha egy két hangminta módosításával a feladat könnyen megvalósítható. Ilyen lehet egyetlen pattanás kijavítása (pl. az elektromos hangszert megérintő játékos elektromos kisülése, egy rossz szájmozdulat a mikrofon előtt stb.), de amint már arról korábban esett szó, egy kiragadott szakasz módosítása után a határok kismértékű korrekciója is. Olyankor is szóba jöhet ez a megoldás, amikor a teljes felvétel csak 1-2 helyen és igen kis mértékben túlvezérlődött (ld. 2.6. ábra kinagyított részét). Ezesetben nem biztos, hogy érdemes nekifutni a teljes felvétel újbóli feljátszásának, sokszor elegendő a túlvezérelt, kinégyszögösedett szakasz széleinek elsimítása – úgy, mintha a jel azon a kis szakaszon dinamika-kompresszáva lett volna.

A szoftveres megvalósítások alig különböznek egymástól, a felhasználónak többnyire teljesen ki kell nagyítania a módosítandó szakaszt (úgy, hogy a hangmintákat egyenként lehessen látni), majd valamilyen rajzoló eszközzel átrajzoljuk a hullámformát (2.6. ábra) avagy a hangmintákat jelölő kis gombócokat megfogjuk és elmozdítjuk (2.7. ábra). Az eredményt természetesen ellenőrizni kell, de a gyakorlott kéz itt is könnyen célhoz ér.



2.6. ábra. Túlvezérelt szakasz kézi javítása „ceruzával”

Bár ez a korrekció nagyon durva beavatkozásnak tűnik, mégis azért nem tekintem a zene meghamisításának, mert csak egy két hangminta átírásáról van szó, ami konkrétan pl. 5 hangminta és 44100 Hz-es mintavételi frekvencia

mellett 0.11 ms időtartam megváltoztatásáról van szó. Ez pedig olyan csekély, hogy a zenei tartalmon ez nem változtat(hat).



2.7. ábra. Hullámforma szerkesztés „kézi tologatással”

## 2.2. Vágási munkálatok

Térjünk vissza egy rövid időre a fejezet elején boncolgatott „meghamisítom avagy sem” témakörhöz. Tekintsük egy kiváló előadó produkcióját, melyet ki szeretnénk adni. Már a felvétel során is előfordulhat, hogy az előadó bizonyos részekben úgy érzi, hogy sikerült ezt már jobban is lejátszani. Így a jobb produkció érdekében újra és újra nekifut a darabnak, de kellő igényesség esetén óhatatlanul érzi, hogy hol ez, hol az a részlet sikerül jobban. Felmerül tehát a kérdés, nem lehetne-e összevágni egy jót? Egyáltalán illik-e, hiszen egy koncert során is csak egyszer adja elő a darabot, úgy ahogyan az sikerül. Igen ám, de a hallgató is csak egyszer hallgatja meg. A kiegyensúlyozottan jó előadásba belebele csúszhat egy-egy „gyengébb” részlet is, mégis jónak értékeljük a darabot, hiszen az élőzene varázsa, a művész kiváló produkciója elfelejteti velünk az apróbb szépséghibákat. Ezzel szemben egy felvételt – különösen, ha szeretjük – sokszor hallgatunk meg, és egy adott helyen előforduló apróbb hiányosságra előbb-utóbb felfigyelünk, és onnantól kezdve zavarni fog minket. Ezért – véleményem szerint – megengedhető, sőt általában szükséges, hogy ne egy egyszeri nekifutást rögzítsünk, hanem több nekifutást összevetve vágjunk össze

egy jó, „megszerkesztett” anyagot. Ez persze nem jelenti azt, hogy ütemenként vagdossunk, mert az óhatatlanul is természetellenesen fog hatni, törekedni kell a minél hosszabb egybefüggő részletekre. Ráadásul az is a megvágás mellett szól, hogy a publikum hiánya, a lehető legjobb felvételi körülmények okozta kellemetlenségek (belógó mikrofonok, visszhangmentes feljátszó, fejhallgató viselete, ingerszegény környezet, stb.) és a hosszan elnyúló felvételi munkálatok nem elhanyagolható mértékben csökkentik az előadás tökéletesre sikerülésének esélyeit.

Ezek fényében tehát célszerű kettéválasztani a „szerkesztés” fogalmat: vágási munkálatokra és utófeldolgozásra. Ezutóbbira gyakran nincs is szükség, mert az előadó játékkészsége, hangszere és a felvételi helység ill. berendezések mind megfelelnek az elérni kívánt hangzásképp eléréshez szükséges feltételeknek. Amennyiben ezek valamelyike nem elérhető, úgy most már kénytelenek vagyunk az eredeti felvétel „meghamisításához” folyamodni. A legegyszerűbb okok közé tartozik, ha nem rendelkezünk jó akusztikai minőségű koncertteremmel és ezért a felvételt valamilyen hangcsillapított környezetben vesszük fel, majd utólagosan adunk térhatást a felvételnek: zengetünk.

Most azonban induljunk ki a szükséges feltételek teljesüléséből és csak azt az esetet vizsgáljuk, hogy az előadó többször játssza fel a darabot. Ilyenkor meg kell állapítani, hogy a darab egy-egy szakasza mely felvételből fog „összeollózódni”, majd ezután össze kell rakni a mozaikot. Kérdés az, hogy lehet-e. A válasz igen – amennyiben a hangzásképp ÉS a játékmód nem változtak. Az előbbi többnyire teljesül, már amennyiben nem változtattunk a felvétel konfigurációján (nem rendeztük át a mikrofonokat – a hangforrás nem távolodott el vagy közeledett a mikrofon(ok)tól, nem cserélgettük őket, nem állítódtak el a felvevőlánc analóg oldalán a szabályozható elemek, stb.). A játékmód meg nem változása már közel sem triviális, különösen a hosszan elnyúló felvételek során az előadó hangulatának, érzetének változásai óhatatlanul is befolyásolják és alakítják a művész játékmódját. Hogy mennyire, az az előadó képességeitől, a nem túl kellemes környezeti tényező kellemetlenségi fokától és a felvétel hosszától függhet. Különösen nehéz a helyzet akkor, ha a felvételt kénytelenek vagyunk egy-egy éjszakai pihenőre megszakítani.

## Crossfade

Ha mindent rendben találtunk és úgy ítéljük meg, hogy két szakaszt össze lehet vágni – vagy legalábbis érdemes megpróbálni, akkor használjuk az átkeverés nevű műveletet. Már az analóg hangtechnika korában is hamar rájöttek, hogy a szalagok egyenes (lejátszási irányra merőleges) elvágása, majd az összeillesztendő darabok összeragasztása nem hoz kielégítő eredményt. A ferdén megvágott és összeillesztett szalagok jobban szólnak. Ennek az az oka, hogy a hirtelen vágásoknál óhatatlanul „megtörik” a hullámforma. Ha viszont ferdén vágunk, akkor az egyenlő azzal, hogy az egyik szakaszt rövid időn belül, de folyamatosan halkítunk el, a másikat pedig hangosítjuk fel. Ha a két felvétel ugyanazon részletére alkalmazzuk a halkulás-hangosodás műveletet, akkor az összevágott jel energiája nem változik, mégis átmegy egyik felvételtől a másikba.

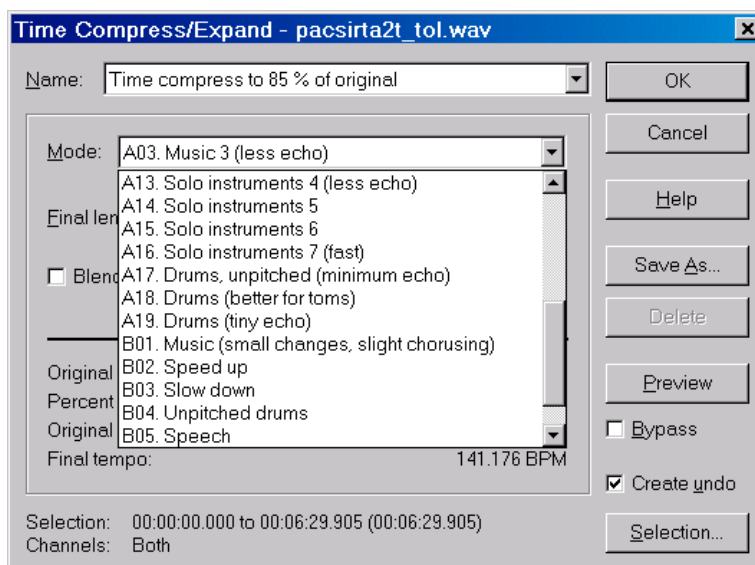
A digitális hangtechnika világában is hasonlóan hajtjuk végre ezt a műveletet, azzal a különbséggel, hogy az átkeverendő szakaszokon nem egyenként hajtjuk végre a fade-out és a fade-in műveleteket, hanem virtuálisan dolgozunk. Kijelöljük, hogy a lejátszandó végtermék milyen felvételek mely darabkáiból tevődik össze. Mindezt úgy, hogy egy-egy váltásnál átfedés van az aktuális két felvétel között. Az átfedésekre pedig – többnyire automatikusan – crossfade-t definiálunk. Meghallgatással ellenőrizzük, hogy jók-e az egyes átmenetek, és ha minden rendben, akkor generáljuk a készre vágott darabot a kijelölések, átfedések alapján.

A crossfade legfontosabb paramétere az átfedés időtartama, azaz, hogy mennyi idő alatt menjen át a zene az egyik felvételtől a másikba. Hasonlóan a fade-in, fade-out műveletekhez, itt is gyakran megadhatjuk a le- ill. felkeverési hangerő-görbe alakját (lineáris/exponenciális).

## Tempó változtatás – time stretch

Bár ez és a következő beavatkozás már talán az effektek című fejezetbe tartoznának, mégis azért ejtünk itt róla szót, mert könnyen előfordulhat, hogy az összevágandó felvételek tempói nem teljesen azonosak. Ez persze metronómmal (taktjellel) készülő felvételeknél nemigen jelentkezik, de ott viszont az előadás természetessége is csorbát szenved (hiszen élőkoncerten sincs taktjel). Ha két

felvétel tempója nem elegendően azonos, akkor mód nyílik a tempóváltoztatásra is. Különböző programok a legkülönbélebb algoritmusokat használják a zene tempójának megváltoztatására. Általánosságban elmondható, hogy kismértékű beavatkozásnak nincsenek feltűnő mellékhatásai, míg nagymértékű lassítás ill. gyorsítás jól hallható ismétlődéses, avagy kórusos hatást eredményez. Univerzálisan optimális algoritmus nincs, a felvétel jellegétől (beszéd, zene, ha zene akkor milyen zene) függ, milyen algoritmusnak hallható a legkevésbé a mellékhatása.



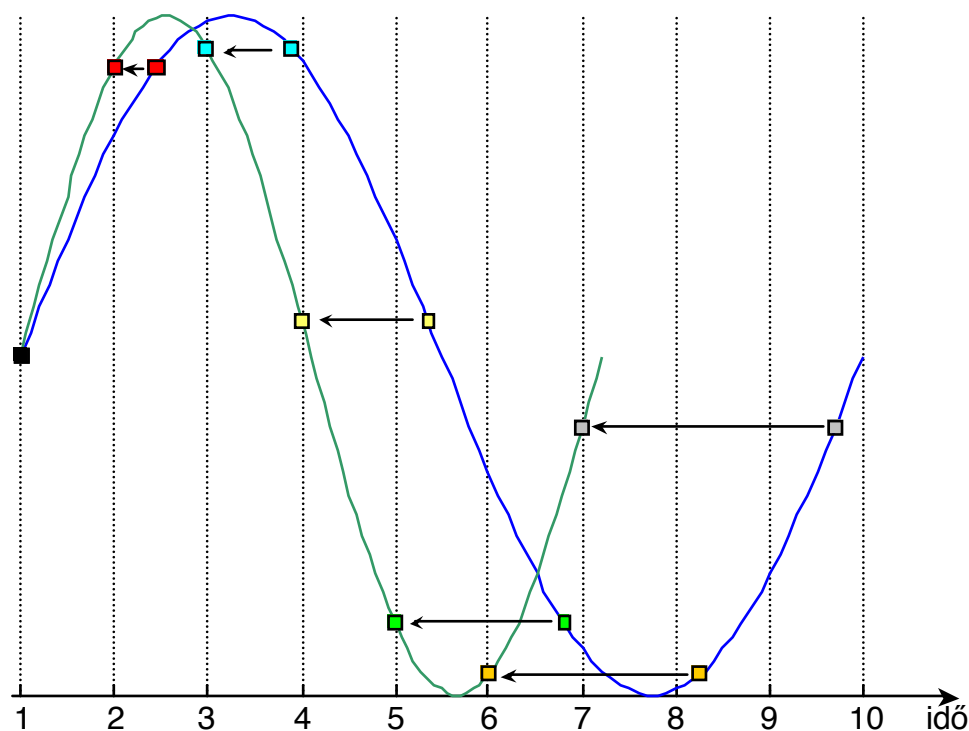
2.8. ábra. Tempó változtatás – széles algoritmusválasztási lehetőség

## Hangmagasság változtatás – pitch correction

Ez a feldolgozási módszer gyakran visszás érzéseket kelt az emberekben: mindenki arra gondol, hogy a tisztán énekelni vagy játszani nem tudó előadó hibáinak leplezésére szolgál. Arra is lehet és szokás használni, de ártalmatlanabb felhasználási terület is adódhat: ha például az egyik hangszert nem jól hangoltuk fel, avagy időközben elhangolódott, és bár önmagához képest tiszta a hangszer, mégis a többi hangszerhez, avagy a korábbi felvételekhez már nem. Ilyenkor néhány cent-nyi hangmagasság módosítás ésszerű alternatívája lehet az újra feljátszásnak.

A hangmagasság módosítás első megközelítésben egyszerű feladat: a felvételt elegendő gyorsabban vagy lassabban lejátszani. Ez a digitális technikára lefordítva azt jelenti, hogy az eredeti (felvételi) mintavételi frekvenciától eltérő

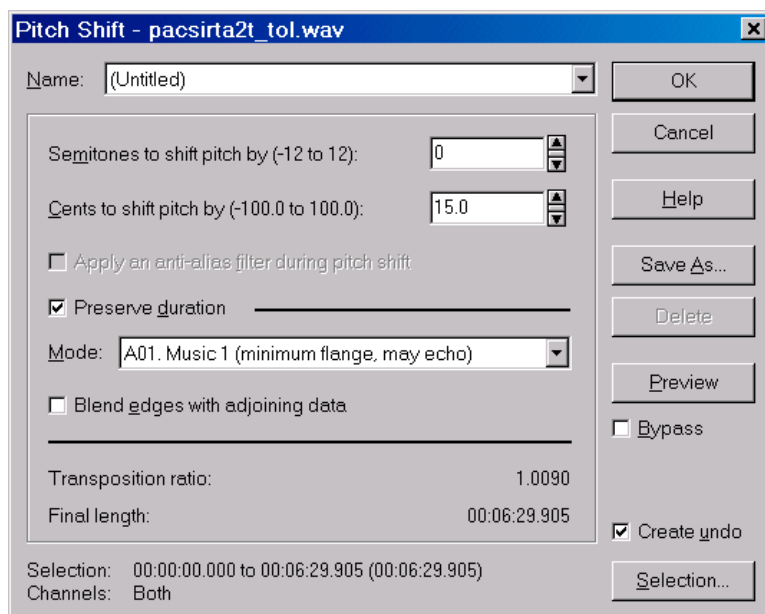
frekvencián játszuk le a számot. Ez azonban általában nem kívánatos (nem is beszélve arról, hogy melyik eszköz játszik le mondjuk 44076 Hz mintavételi frekvencián), ezért az eredeti jelet újra kell mintavételezni, és ezután mégis az eredeti mintavételi frekvenciával lejátszani (ld. 2.9. ábra).



**2.9. ábra. Újramintavételezett jel lejátszása az eredeti frekvencián**

A késsel jelölt görbe az eredeti jel, a mintavételi időpontok az időtengelyen megjelölt 1-10 pontokban vannak. Az új mintavételi pontok a kék görbén jelölt négyzetek, melyeket azután visszatolunk az eredeti mintavételi időpontokba és így kapjuk a zöld görbét.

Ha alacsonyabb frekvencián mintavételezzük újra a jelet, akkor magasabb hangot kapunk. Az alacsonyabb mintavételi frekvenciához azonban alacsonyabb maximális frekvencia tartozik, ezért visszahajlási jelenségek elkerülése érdekében szükség lehet előzetes aluláteresztő szűrőre (ld. 2.10. ábra „Apply an anti-alias filter during pitch shift” opciót). Gyakoribb azonban, hogy a pl. magasabban lejátszással együtt járó tempógyorsulás nem engedhető meg, ezért többnyire lehetőség nyílik az eredeti tempó visszaállításához szükséges korrekció kiválasztására is és az előző bekezdésben említett tempóváltoztatási algoritmusok egyikét kiválaszthatjuk.



2.10. ábra. Hangmagasság változtatás

## 2.3. Effektek

Aki kizárólag klasszikus és hamisítatlan felvételekkel kíván foglalkozni, az ezt a fejezetet nyugodt lélekkel átugorhatja. Itt ugyanis a könnyűzenében mindig, de klasszikus zenében is időről-időre alkalmazott beavatkozásokról van szó. Itt ténylegesen az a cél, hogy a zenét megmászítsuk, kellemesebbé, élvezetesebbé, hallgathatóbbá tegyük. Felhasználási súlyuknak megfelelően ebben a fejezetben csak a leggyakoribb effektek kerülnek nagyító alá. Ezeken felül még rengeteg létezik, korlátot csak az emberi fantázia szab. Az általános szabály, hogy próbáljuk ki, hallgassuk meg, és ha tetszik, használjuk.

### Hangszínszabályozó – EQ

Ezt talán senkinek sem kell magyarázni, hogy miről is van szó, hiszen mindenki látott már a különböző Hi-fi tornyokon, Music centereken hangszínszabályzó gombokat. Célként két dolgot említenék meg: egyrészt a felvételi eszközök ill. a hangforrás hibáinak kijavítását, másrészt a „dögösebb” hangzások elérését. Az előbbi kategóriába tartozik pl. a gyengébb minőségű mikrofonok frekvenciakarakterisztikájának korrekciója, az utóbbiba pedig mondjuk lábdob basszusának kiemelésével elért átütőbb hangzás. Ezek általában kisebb mértékű hangszínezést jelentenek, de használhatjuk szélsőségesen pl. egy adott frekvenciatartomány kiszűrésére is.

A komolyabb szoftverekben található hangszínszabályozók a hagyományos grafikus hangszínszabályzók mellett a parametrikus hangszínszabályozást is támogatják. A leggyakoribb paraméterek a következők:

- a vágási karakterisztika (alul-, felül-, sáváteresztő ill. sávzáró szűrők)
- középfrekvencia (ahol a legnagyobb a hatás)
- jósági tényező (alul-, felüláteresztő szűrőknél a meredekséget, sáváteresztő, -záró szűrőknél pedig a tartomány szélességét befolyásolja)
- emelés/vágás mértéke (a középfrekvenciánál alkalmazott érték decibelben)



Ezen paraméterek mellett a következőkkel is találkozhatunk:

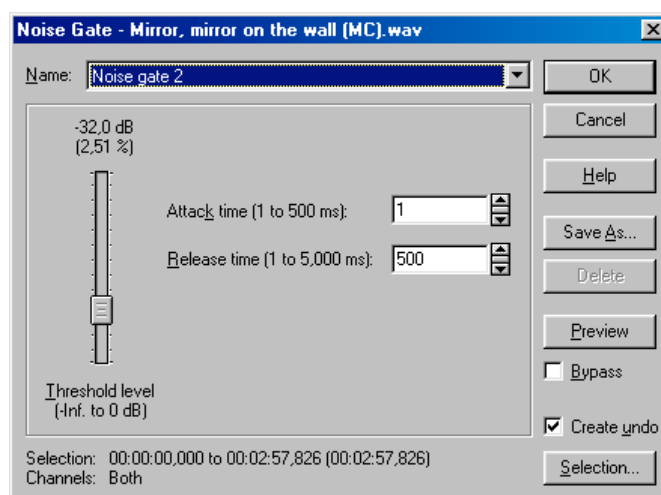
- alkalmazott szűrő típusa (Butterfly, Chebisev, stb.)
- a választott szűrő fokszáma (ez határozza meg a maximális meredekséget)
- szűrési tartomány szélességének megadása – oktáv, terc, 1/8 oktáv szélesség (ez a jósági tényező paramétert helyettesíti, a zenész gondolkodásmódhoz közelebb álló felfogással)

Hasonlóan analóg felvételekhez, az utólagos hangszínszabályozásnak mindig megvan az a hátránya, hogy a hanghordozó zaja (szalagzaj analóg esetben, digitális esetben kvantálási zaj) a jellel együtt erősödik, és ha egy eredetileg halk frekvenciatartományt jelentősen kiemelünk, úgy a korábban nem meghallható zaj esetleg zavaróan hallhatóvá válik. Itt szintén szélsőséges esetekről van szó, de ha biztosan tudjuk, hogy korrekciót kell végrehajtanunk, akkor erről már a hangrögzítéskor, az analóg oldalon gondoskodhatunk – már amennyiben azon eszközeink jel/zaj viszonya jobb (a módosítandó frekvenciatartományban), mint a jelerősségből és a számábrázolásból adódó kvantálási zaj viszonya. Érdeemes megjegyezni, hogy a pszichoakusztikus kódolású felvételek egyes sávokban különösen rossz jel/zaj viszonyúak lehetnek annak ellenére, hogy nem halljuk meg. Néhány dB kiemelés azonban megemelheti ezt a meghallás küszöbe fölé és ez komoly gondot okozhat. Ezért még feldolgozandó anyagot nem célszerű pszichoakusztikus kódolással tömöríteni, vagy ha mindenképpen szükséges, akkor a lehető legkisebb mértékben tömörítsünk.

### **Zajzár – (noise) gate**

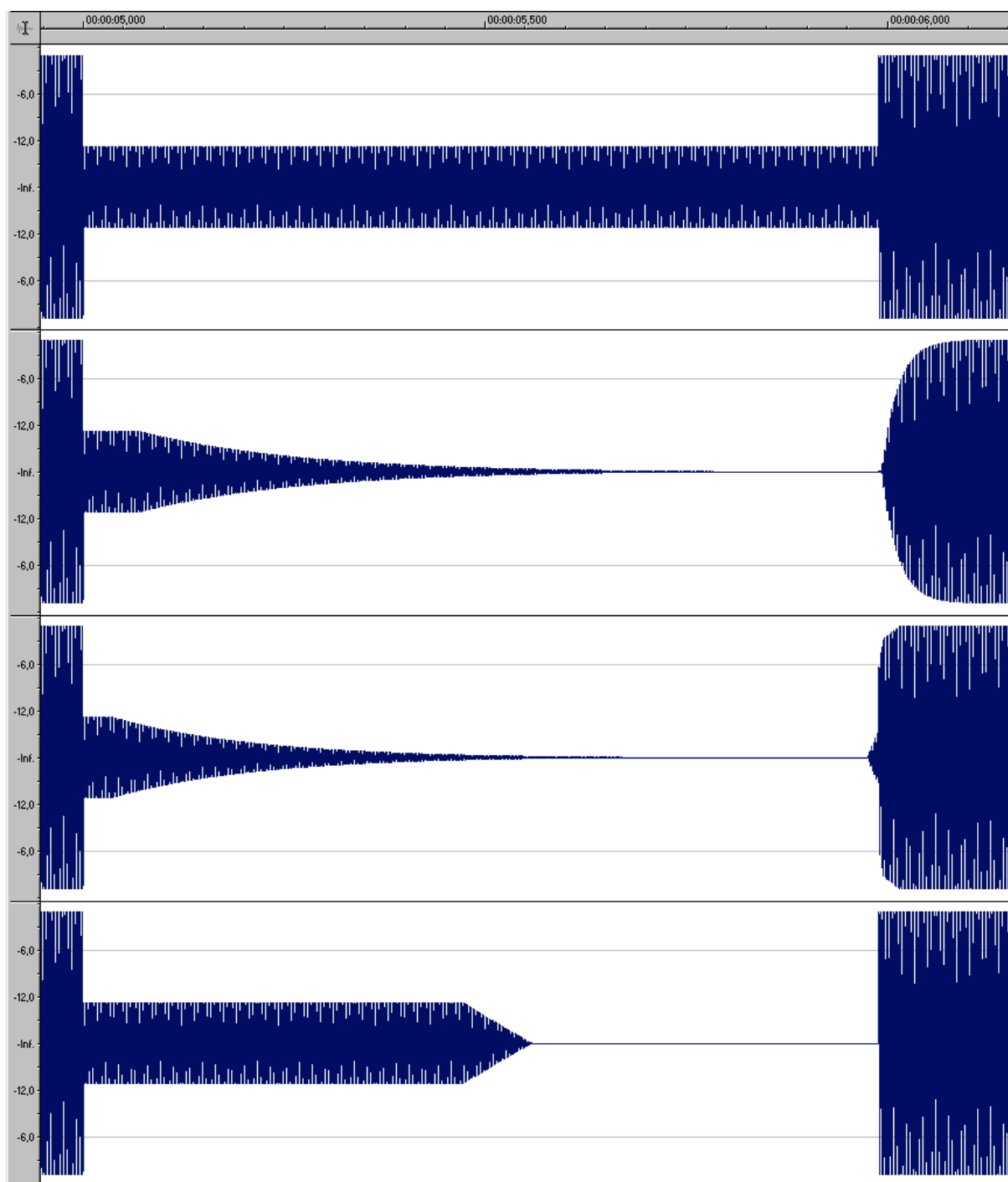
Az effektek között szokás beszélni róla, mivel alapvető eszköz a felvételkészítés és előprodukció során is, holott hatását tekintve inkább a zajcsökkentés fejezetbe tartozna. A szokásos szemléletmód megtartásához beszéljünk tehát itt róla, annál is inkább, hogy néhány paraméterének megismerése és megértése segítséget nyújthat a kompresszor működésének megértésében is.

A klasszikus zajzár nem más, mint egy olyan eszköz, mely adott bemenő jelszint alatt (küszöbszint) a kimenetet némítja, felette pedig egyenesen átengedi. Így ha tudjuk, hogy a bemeneti rendszer alapzaja pl. -60 dB-nél van, a jelszint pedig lényegesen e felett, akkor nyugodtan alkalmazhatunk egy olyan zajzárat, mely -50 dB alatt némítja a kimenetet; azaz ha a jel megszűnik, és csak alapzaj érkezik be, úgy azt kinémítja ez a berendezés.



2.11. ábra. A zajzár

A küszöbszint megadásán túl a vezérlést két időparaméterrel adjuk meg: a megfogási (attack time) és az elengedési (release time) idők beállításával (2.11. ábra). Ezek a zár beállításának (elengedési idő) ill. feloldásának (megfogási idő) hosszát határozzák meg. A természetes hangzás érdekében ui. nem némítanak ugrásszerűen, hanem a jel küszöbszint alá csökkenése után elengedési paraméterrel megadott időtartam alatt folyamatosan némul el a kimenet. (Ennyi ideig tart a „hangerő-szabályozó lecsavarása”.) A megfogási paraméter pedig azt adja meg, hogy ismételt jelérzékelés esetén (a bemenő jel meghaladja a küszöbszintet) mennyi idő alatt hangosodjon fel a jel a némáról a bemenő jel szintjére. (Megjegyzés a terminológiával kapcsolatban: Az elnevezések a kompresszoroknál (ld. következő effekt) használt elnevezéseket követik, de mivel ott pont a küszöb fölötti jelre kell beavatkozni, míg itt a küszöb alattira, ezért ezeknek a paramétereknek az elnevezése pont fordított a zajzár hatásmechanizmusával.) Ezen átmenetek megvalósítása nagyon különböző lehet (2.12. ábra).



**2.12. ábra. Zajzár megfogási és elengedési idejének különböző megvalósításai**

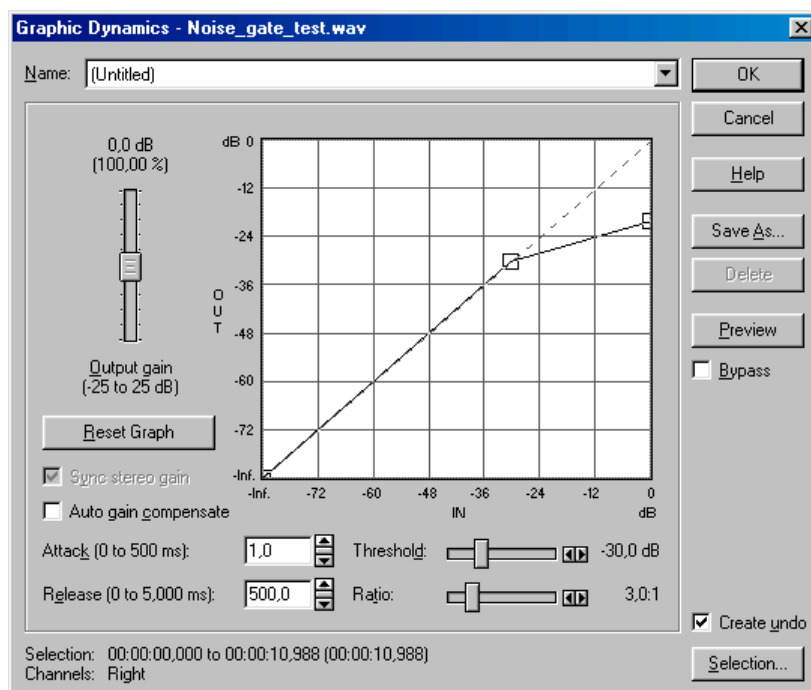
Késleltető nélküli analóg eszközök csak azután tudnak reagálni, hogy az esemény megtörténik, tehát a zajzár fölengedési ideje miatt a megszólaló hang elejéből valamennyit levág a zajzár. Digitális megvalósítás esetén ez nem gond, mivel általában úgy működnek ezek az effektek, hogy egy bizonyos darabot letárolnak, azt analizálják, feldolgozzák és csak ezután engedik ki. Ez általában még valós idejű feldolgozásnál sem gond, mivel néhány hangmintányi (0.1 ms-os

nagyságrendű) késleltetés nem hallható meg. Változhat ezen felül a lehalkítási ill. visszahangosítási görbék alakja: lineáris vagy exponenciális, szélsőséges esetben lépcsős.

Komolyabb eszközök esetében a zajzár az nem feltétlenül zár, hanem csak adott (szabályozható mértékű) hangerővel halkít. Ezt párosítva azzal, hogy egy csatornát analizálva megkapjuk, hogy hol kell halkítani, és azt a halkítást egy másik csatornán hajtjuk végre olyan eszközhöz jutunk, melyet a rádiózás során folyamatosan használnak. Ducking-nak hívják, és pl. arra jó, hogy a bemondó megszólalásakor a zene automatikusan lehalkul.

### **Dinamika-kompresszor – compressor, dynamics processor**

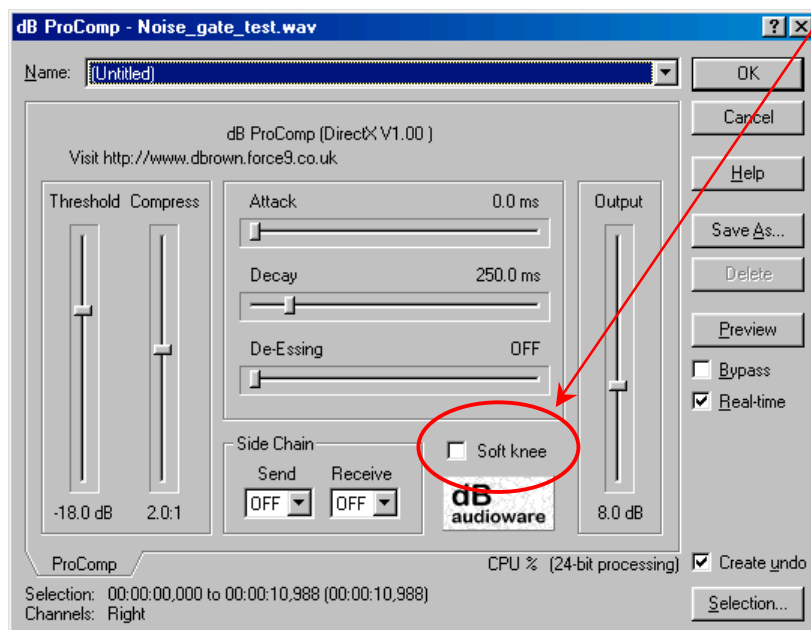
Amint azt a neve is mutatja, olyan módszerről van szó, mely a jel eredeti dinamikatarományát csökkenti. Többsávós felvételeknél a sávokat egyenként, majd a kész kevert jelet is szokás kompresszoron átvezetni. Általában azért használják, hogy a jel/zaj viszony nőjön. Itt azonban nem elsősorban a lejátszó berendezések zajáról van szó, hanem pl. közúti zajról: egy adott hangerő fölé nem akarunk menni, hogy a figyelmünket ne vonja el a vezetéstől, viszont ha túl halk a zene, akkor nem haljuk a motorzajtól, a futómű zajától, a macskakótól stb. Mivel pl. autóban ez a hangerőtartomány jelentősen kisebb, mint egy élő előadásé, ezért autóban a dinamikatarományban nem korlátozott felvételek hallgathatatlanok. Egy másik példa a jel/zaj viszonyra az énekes érthetősége. A szöveg megértése szempontjából a kísérő zene zajnak minősül, tehát ha az ének túlságosan elhalkul a kísérethez képest, akkor a szöveg érthetlenné válik, és ez a felvétel minőségét általában negatívan befolyásolja. A kompresszor speciális esetben különleges hanghatások elérésére is használható, ez hamarosan bemutatásra kerülő paraméterek megfelelő beállításával érhető el.



**2.13. ábra. Kompresszor grafikus átviteli görbével**

Egy kompresszor legfontosabb paramétere az átviteli görbéje. Erről logaritmikus léptékben (azaz dB-ben) le lehet olvasni, hogy adott bemeneti szinthez milyen kimeneti szint tartozik. A 2.13. ábrán a görbe két szakaszt tartalmaz: egy 45 fokosat, mely -90 dB és -30 dB között 1:1-es átvitelt eredményez (azaz -30 dB-nél alacsonyabb jelszint esetén nincs hangerő-beavatkozás), felette azonban a bemeneti 30 dB-es tartomány (-30 és 0 között) 10 dB-es kimeneti tartományba van tömörítve (-30 és -20) közé. Ez azt jelenti, hogy a tömörítési arány 3:1, mivel 30 dB bemeneti tartományhoz 10 dB-es kimeneti tartomány párosul. A görbe töréspontját – ahonnan kezdve hatásos a kompresszor (jelen esetben a -30 dB-es pont) – küszöbszintnek nevezik. Az angol megfelelők: Threshold (küszöbszint) és Ratio (tömörítési arány). A görbéről az is leolvasható, hogy a maximális kimeneti jel -20 dB, tehát (ha az átviteli görbét minden egyes hangmintára egyenként alkalmazzuk, akkor) a teljes jelszint 20 dB-el emelhető. Ennyi a tömörítés nyeresége lehallgatás során változatlan külső zajviszonyok mellett. Egyszerűbb kompresszorokon csak ezzel a három paraméterrel (küszöb, arány és hangerő-korrekciónak) lehet megadni a kompresszor átviteli görbét. Grafikus kezelőfelület esetén azonban szinte mindig lehetőség nyílik tetszőleges karakterisztikák kialakítására és több töréspontot definiálhatunk. Ezáltal pl. simább átmenetet képezhetünk az egyes jelszintek

között. Sokszor még arra is lehetőség nyílik, hogy az egyenesekből összerakott görbe helyett egy simított görbével dolgozzon a kompresszor (2.14 – soft knee).

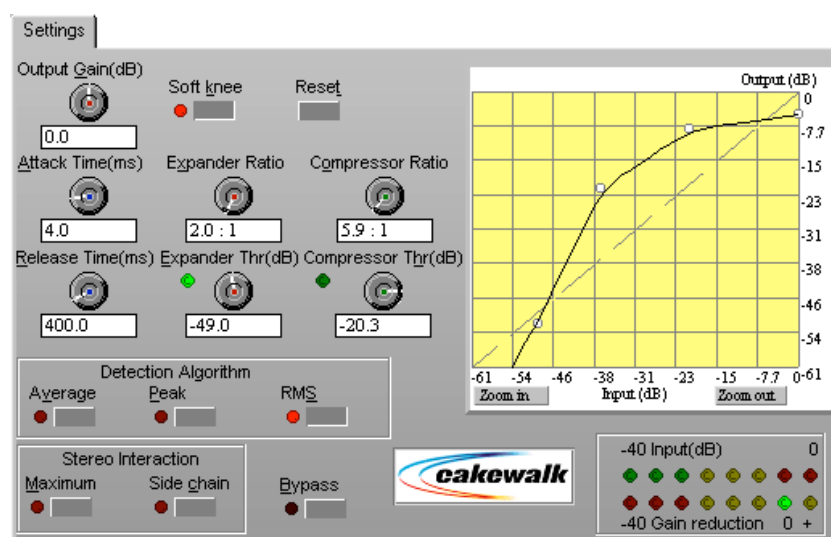


**2.14. ábra. Kompresszor simított átviteli görbével a töréspont körül (Soft knee)**

A zajzárhoz hasonlóan a kompresszor működése sem hangmintáról-hangmintára való megvalósítás, hanem itt is van megfogási és elengedési idő. Ez azonban lényegesen bonyolultabb megvalósítást igényel, mint a zajzáré. A legegyszerűbben úgy képzelhető el, hogy hangmintáról-hangmintára végigmenve a kompresszor meghatároz egy erősítési tényezőt a kompresszor átviteli görbéje alapján (dB-ben). Majd az így kapott erősítésvezérlési görbét, mely sok meredek ugrást is tartalmazhat, megsimítja a vezérlő algoritmus, hogy ne legyenek annyira hirtelen erősítésváltozások. Ez még annyiban bonyolódik, hogy a lefelé ill. felfelé történő erősítés-változtatásokhoz különböző időállandójú simítást kell alkalmazni a megadott megfogási- és elengedési idők alapján. Az algoritmus nehézsége abban rejlik, hogy a fel- és lefutási idők folyamatosan egymásba nyúlnak... Amikor előállt a logaritmikus erősítésvezérlési görbe, akkor ezt át kell váltani lineárisra, majd az így kapott görbével hangmintáról-hangmintára végigszorozni a jelgörbét.

Az igazsághoz hozzátartozik, hogy az imént vázolt megvalósítás csak egy lehetséges megoldás, ami már első közelítésben is változni fog, ahhoz, hogy a kompresszor valós-idejű effektként is használható legyen. Ilyenkor a zenének

mindig csak egy rövidebb darabkája áll rendelkezésünkre, ezen kell a módosításokat a lehető legkisebb késleltetéssel végrehajtani. Egy másik szempont, hogy a jobb hangzás érdekében a megfogási- és elengedési idők definiálása igen szabadon értelmeződik. Eredetileg ui. tiszta volt a kép: ha a küszöbszintet alulról átlépjük, akkor megfogási idő van – hiszen a küszöbszintet meghaladó jelet „vissza kell fogni”. A másik esetben (küszöbszint felülről átlépése) pedig értelemszerűen az elengedési időt kell alkalmazni. A szabadon megadható átviteli görbékkel azonban a küszöbszint fogalma nem egyértelmű, hiszen olyan görbét is lehet definiálni, aminek az alja nem 45 fokos meredekségű, azaz nincs olyan (hosszabb) szakasz, ahol az átvitel 1:1-es lenne (2.15. ábra). Innentől kezdve pedig mindenki azt valósít meg, amit tud és amit a legjobb hangzásúnak hisz. Valósídejű effekteknel pedig az is jelentős szempont, hogy mekkora az adott algoritmus számításigénye, ezért sokszor nem is törekednek a lehető legjobbra, hanem egy egyszerűbb, de minimális processzorigényű algoritmus megvalósítása mellett döntenek. Ily módon az amplitúdó-változtatás karakterisztikája is a fülünkhöz legközelebb álló logaritmikus karakterisztika helyett lehet lineáris, avagy lépcsős is.



2.15. ábra. Kompresszor karakterisztika 45 fokos szakasz nélkül

Összegzésként elmondható tehát, hogy a megfogási- és elengedési idő valamilyen módon a kompresszor beavatkozási sebességét határozzák meg, de ezek az értékek mindenképpen csak körülbelülinek tekintendők. Itt is az

érvényes, hogy állítsunk be egy tervezett értéket, és meghallgatás után módosítsunk, ha szükséges.

Érdemes szót ejteni az analóg kompresszorokhoz képesti lényeges különbségről is: a digitális hangminták tárolása miatt a kompresszort megvalósító algoritmus képes lehet 0 idő alatt reagálni, azaz a felmenő csúcsot egyből megfogni. Ezzel szemben az analóg eszközök mindenképpen csak minimális késleltetéssel tudnak beavatkozni, így a túl meredek felfutású „beütések” az átlagos jelszinten jócskán túllőhetnek, és ez a későbbi digitalizálás szempontjában csökkenti a zene hasznos dinamikatartományát. Ily módon, ha a digitális kompresszor 0-nál nagyobb megfogási időt használ, akkor a gyorsan felfutó csúcsok helyben maradhatnak, és nem igaz az, hogy a korábban példaként bemutatott 3:1-es tömörítésű, -30 dB küszöbszintű kompresszor tiszta nyeresége 20 dB, minél nagyobb a megfogási idő, annál jobban „túl tud lőni” a jel a tervezett görbéhez képest és ezzel mindenképpen számolni kell. Nem valós idejű feldolgozás esetén bátran készítsünk egy próbaállományt hangerő-kompenzáció nélkül, elemezzük a hangállományt és a legnagyobb csúcshoz képesti maximális erősítést adjuk meg hangerő-kompenzációra a kompresszornak, majd számoltassuk újra a teljes hangállományt. (Ne a már letárolt hangállomány hangerejét változtassuk, mert így nagyobb lesz a kvantálási zaj mértéke. Az algoritmusok többsége ui. feldolgozás során lényegesen nagyobb számolási pontossággal dolgozik, mint a letárolt hangminták pontossága.)

## Limitér

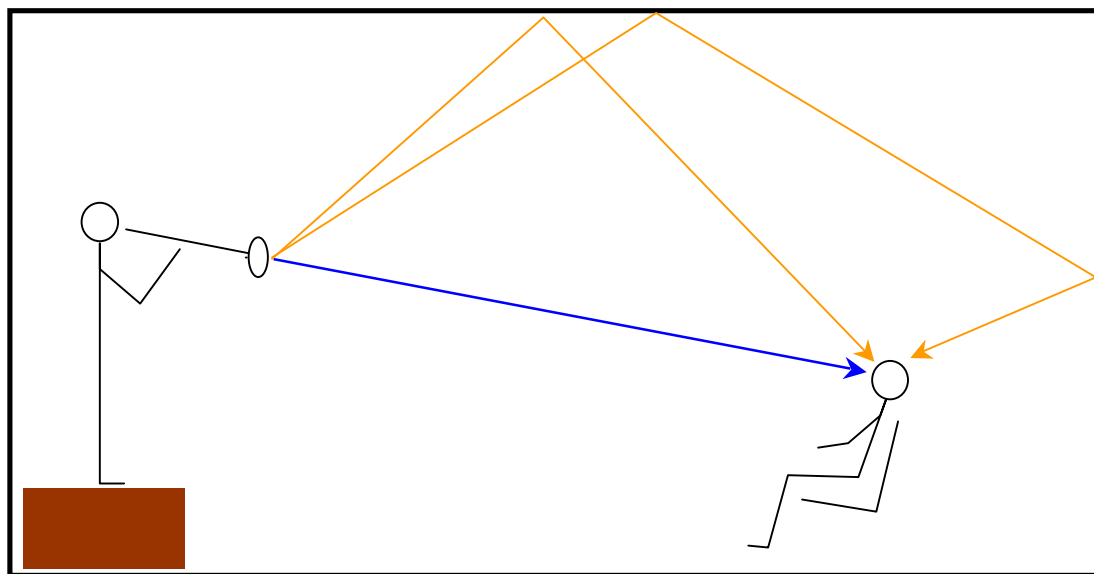
Ha egy töréspontos kompresszor tömörítési arányát nagyon nagyra ( $\infty$ -re) vesszük, és a megfogási időt amennyire lehet a 0-hoz közelítjük, úgy a jel nem képes egy bizonyos szint fölé menni. Ezt az effektet általában a hangszórók védelmére szokás használni, de speciális hangzása miatt használhatjuk céleffektként is.

## Zengető – reverb

Megszoktuk, hogy élő produkció többnyire nagy méretű és jó hangzású termekben hangzik el. Ezzel szemben a felvételek lehallgatása többnyire kisméretű helységekből történik, melyek hangzásképe teljesen más. A



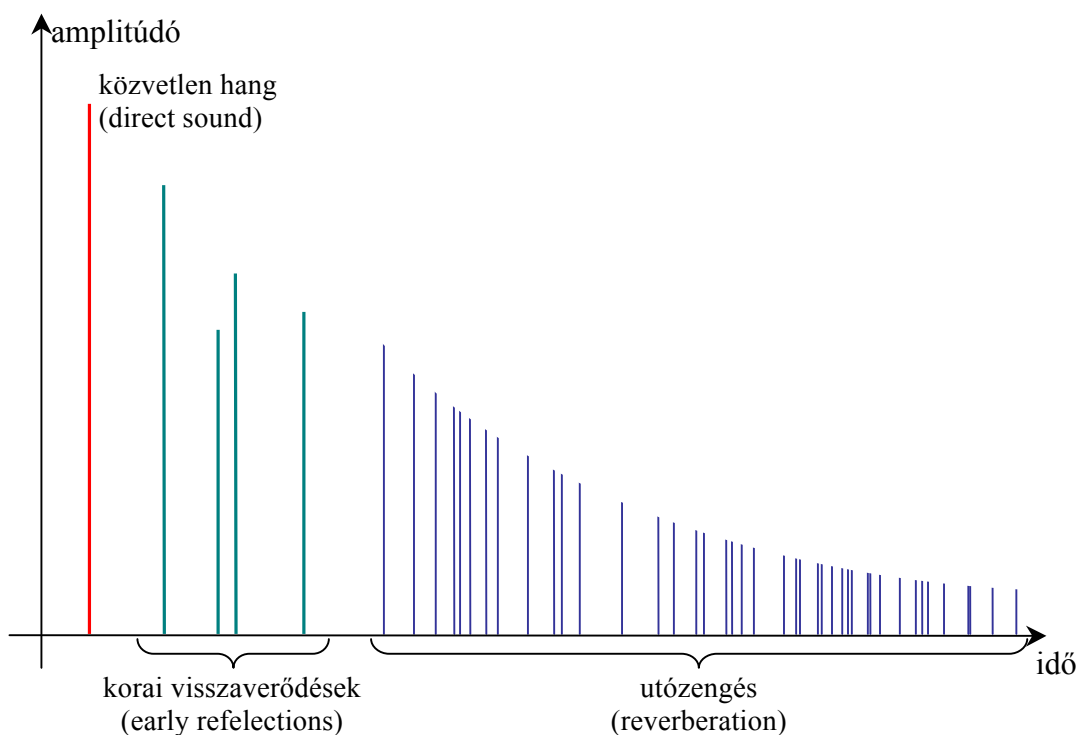
legnagyobb különbség leginkább úgy fogalmazható meg, hogy a „térhatás otthon kicsi”. Ez műszakilag azt jelenti, hogy a sok elnyelő felület (szoba bútorzata) és a rövid visszaverésmentes hangutak miatt az átlagos lehallgató helyiségek (pl. nappali) utózungési ideje sokkal kisebb, mint azé a teremé, ahol amúgy az élőprodukció elhangzana. Hogy mit lehet ez ellen tenni, ahhoz először ismerjük meg az utózungés folyamatát.



**2.16. ábra. Teremreflexiók és a közvetlen hang**

Képzeljünk el egy nagy méretű termet, ahol valahol a színpadon áll egy trombitás, mi pedig valahol a terem másik felében hallgatjuk. Mikor a zenész belefúj a hangszerébe, akkor legelőször a közvetlen hang éri el fülünket (2.16. ábra – kék görbe). Nagyon kis késleltetéssel érkeznek be az első visszaverődések a két oldalfalról és a mennyezetről, hiszen alig hosszabb utat tesznek meg. (A talajról kevésbé, mert ott sok az elnyelő felület: székek, hallgatóság). Ezek a visszaverődések egy idődiagramon ábrázolva még jól elkülönülnek, és a teremérzet szempontjából különösen sok információt hordoznak. Az „idő múlásával” beérkeznek a többi reflektált hangok is, de ezek a különböző kb. azonos hosszúságú utak nagyon sokféle megvalósítása miatt (a pontos útvonal rendkívül sokféle lehet), egyre inkább egymásba mosódnak. Az egyre több és több visszaverődés miatt az amplitúdók egyre kisebbek, és így a jel szép lassan lecseng. A 2.17. ábrán egy terem ún. echogramját láthatjuk. Az első „csík” jelöli a közvetlen hangot, amely megszólaláshoz képesti késleltetése (0 időpillanat)

közvetlenül számolható a forrás és a hallgató távolságából (késleltetés=távolság/hangsebesség). Az ezt követő szakaszt korai visszaverődéseknek nevezzük (zöld csíkok), végül a már egymásba mosódó szakaszt lecsengési szakasznak.



**2.17. ábra. Terem visszaverődések idődiagramja – az ún. echogram**

Ha egy teremben egy statikus hangforrás bekapcsolása után megvárjuk az állandósult állapotot (amikor már nem növekszik a hangnyomásszint adott pontban), majd ezután a forrást kikapcsoljuk, akkor adott pontban a teremben az idő előrehaladtával mérhető a hangenergia folyamatos csökkenése. Azt az időt, ami a kikapcsolástól addig telik el, míg a hangenergia 60 dB-el alacsonyabb szintű lesz, mint az állandósult állapotban volt, utószengési időnek hívjuk. Ez meghatározó paramétere minden zenetűnek.

Mivel nagy termekben hosszabbak a visszaverődés-mentes szakaszai a hangutaknak, ezért a lecsengés többnyire hosszabb, mint az otthoni lehallgató helyiségekben. Így a „zengetés” fogalma azt jelenti, hogy többlet visszhangokat adunk a felvételhez, hogy a lehallgatás során lehetőleg ugyanazokat a visszhangokat (is) halljuk, amiket a koncerten. Ha egy jó minőségű

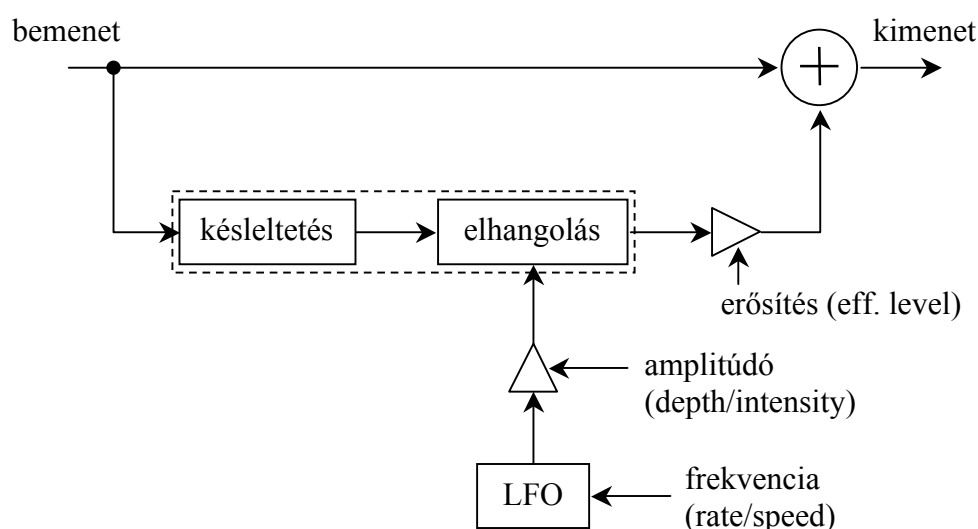
koncerttermet tökéletesen le tudnánk utánozni, akkor az ideális eset az lenne, ha csak a közvetlen hangot rögzítenénk felvételkedzéskor (visszhangmentes környezet), és a lehallgatás is visszhangmentes helyen történne, hogy további, „hamis” visszhangok ne jöjjenek a jelhez. Mivel ez utóbbi nemigen valósítható meg (általában nem süketszobában hallgatunk zenét), ezért nem is cél a tökéletes zengető megvalósítása, hanem olyat kell készíteni, ami egy referencia lehallgató-helyiség zengésével együtt közelíti meg nagyon jól a koncerttermet. A felvételt viszont relatív „száraz”, kevésbé visszhangos környezetben készítjük, hogy minél kevesebb „hamis” visszhang kerüljön a felvételbe. (Azért nem teljesen visszhangmentes környezetet használunk felvételkedzés során, mert az a zenészek komfortérzetét jelentősen lecsökkentené.) Ez a folyamat az általános könnyűzene esetén, klasszikus felvételeknél azonban sokszor a felvételt olyan helyen készítik, ahol amúgy is elhangzana a produkció, és mivel a hallgatóság hiánya a terem akusztikáját befolyásolja, ezért a mikrofonok megfelelő pozicionálásával igyekszenek olyan felvétel készíteni, amely természetesen hangzik. Általános, hogy minimum egy mikrofont helyeznek a hangforrás(ok)hoz közel – ez veszi a direkt jelet, és minimum egyet a távolytérben, amely kifejezetten a terem zengését hivatottak megfogni.

Mivel a könnyűzenei felvételeket tipikusan stúdiókban rögzítik, ill. jó koncerttermet bérelni nem olcsó, ezért legtöbbször közeli mikrofonozással ill. vonalból történő bejátszással rögzítik a zenét. Az így készült felvételek rendkívül szárazak, így feltétlenül zengetni kell őket. Ezért a zengető az egyik leggyakrabban használt effekt. A megvalósításra garadával vannak ötletek, hogy kinek milyen sikerült készítenie, az sokszor szerencsén is múlik. Itt csakis az igaz, hogy a nagy avagy kicsi terem eldöntésén túl semmit nem tudunk meghatározni, és kénytelenek vagyunk addig keresgélgni, amíg nem találunk olyan hangzást, ami tetszik.

A különböző zengetők és visszhangosítók kezelőfelülete rendkívül sokféle lehet, ízelítőül lapozzunk a 3-as fejezet végére (3.19. és 3.20. ábrák).

## Kórus – chorus

Egy kórus attól kórus, hogy sok ember vagy hangszer játssza ugyan azt, de nem pontosan ugyanolyan tempóban, hangszínben, hangmagasságon és nem ugyanazon pontban. A kórus effektnek tehát az a dolga, hogy az eredeti jelet „kissé” elhangolja, késleltesse majd az eredeti jelhez hozzáadja. A viszonylag nagy (10-30 ms) késleltetésre azért van szükség, hogy kellően nagy legyen a kórushatás (ha az egyes kórustagok elég messze ülnek egymástól, akkor a hangutak hossza a kórus egyes tagjaitól a hallgatóhoz is jelentősen különböző), ill. mert nélküle az – amplitúdóban történő összegzés miatt kialakuló – eredő fésűszűrő hatás relatív távoli beszakadásai jól érzékelhető hangszínmódosulást eredményeznének. A konkrét megvalósítás itt is sokféle lehet, a 2.18. ábrán egy lehetséges megoldást látunk. A megadható paraméterek leginkább az elhangolás sebessége (kb. 1-5 Hz) és mélysége (mértéke), ill. a módosított jel eredetihez képesti szintje.



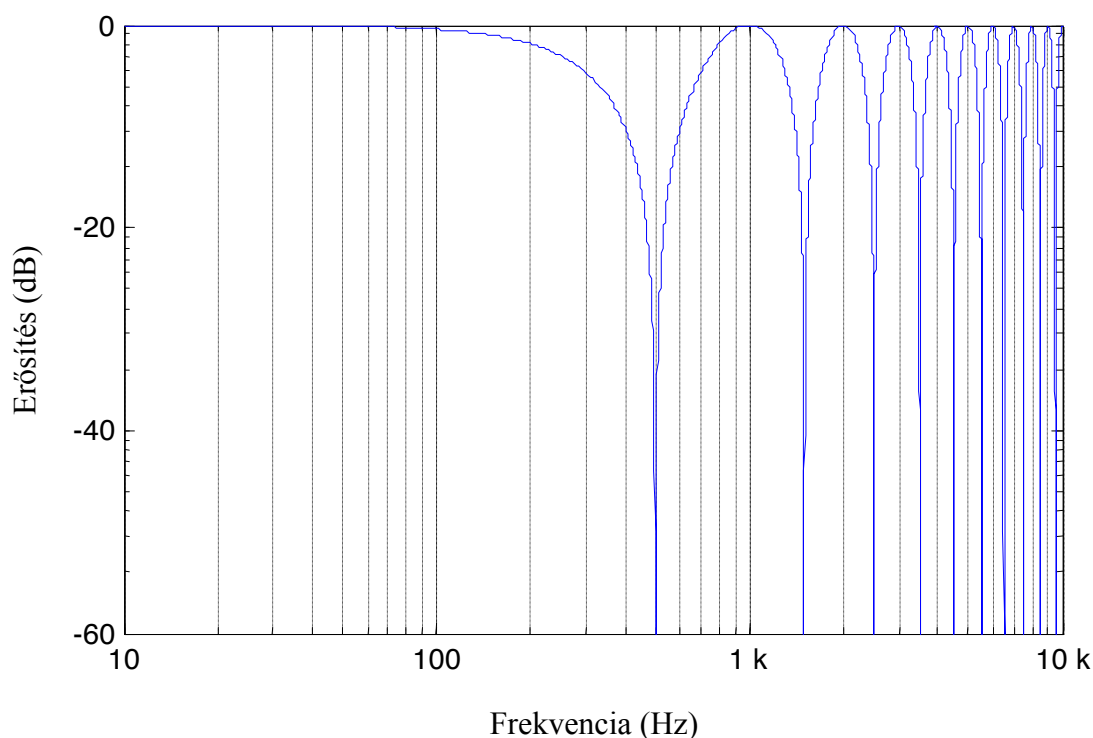
**2.18. ábra. Kórus effekt funkcionális rajza**

Az kisfrekvenciás oszcillátor (LFO) frekvenciájának megadásával, majd az oszcillátor által kiadott (pl. szinuszos) jel amplitúdójának beállításával kaphatunk egy olyan vezérlőjelet, mely az elhangolás sebességét (LFO frekvencia) és mélységét (LFO jelének amplitúdója) egyaránt vezérli. Az elhangolás egyszerű gyorsabb és lassabb lejátszást jelent; ha a vezérlő jel szimmetrikus, akkor átlagosan nem lesz se gyorsabb, se lassabb a lejátszás, viszont a lokális gyorsítás és lassítás folyamatos hangmagasság-változtatáshoz vezet.

## Lebegtető – flanger

Az első flanger-t úgy hozták létre, hogy a zenei anyagot 2 darab szalagos magnetofonra rögzítettek egyidőben, majd újbóli lejátszáskor a két jelet összegezve küldték a hangszórókra. A két magnetofon kismértékű sebességeltéréseinek hatására szólalt meg a jellegzetes lebegő hangzás. Az effekt hatásának növelésére az egyik magnó lejátszási sebességét mechanikusan (fékezéssel) időnként csökkentették, majd amikor elengedték, akkor újra visszagyorsult a magnó. Így valósították meg a folyamatosan változó, markáns lebegő hangzást.

Ha a fenti módszert jelfeldolgozás szempontjából vizsgálunk, akkor a következőket állapíthatjuk meg: ha az eredeti jelhez egész rövid idővel késleltetett változatát hozzáadjuk, akkor az nem más, mintha a jelet egy fésűszűrőn eresztenénk át, ahol a 0 dB-es pontok frekvenciái fordítottan arányosak a késleltetési idővel, a leszívási frekvenciák pedig pontosan közöttük vannak (2.19. ábra).



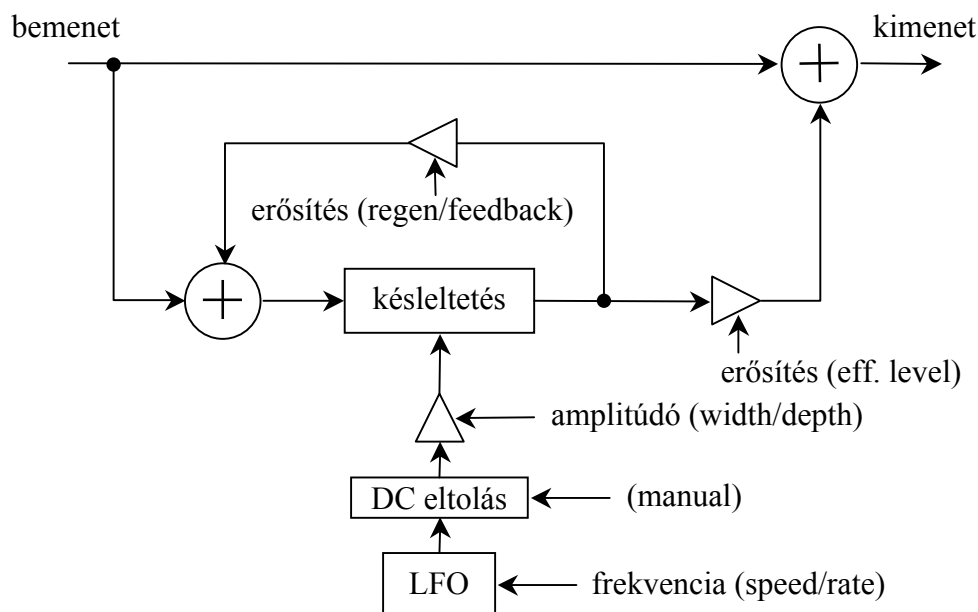
2.19. ábra. Flanger átviteli görbéje 1ms-os modulálatlan késleltetési idő esetén  
(visszacsatolásmentesen)

Ha a késleltető időparaméterét kis mértékben, folyamatosan, de lassan változtatjuk (egy kisfrekvenciás oszcillátorral moduláljuk), akkor máris elkészítettük a flanger-nek nevezett effektet. Ez a frekvencia-átvitel szempontjából annyit jelent, hogy a leszívási pontok szép lassan „vándorolnak” ide-oda.

Vegyük észre, hogy megvalósítás szempontból ez teljesen ugyanaz, mint amit a kórusnál is alkalmaznak. Az ott leírt elhangoló ui. nem más, mint egy gyorsan változó idejű késleltető, amely egy szimmetrikus plusz-mínusz tartomány között változik. Ha az „elhangoló” előtt található késleltetővel ezt összevonjuk, akkor az azt jelenti, hogy egy adott (relatív nagy) késleltetési idő körül változtatjuk kis mértékben a tényleges késleltetést.

A lényeges különbség abban rejlik, hogy a kórus késleltetési ideje min. 10-30 ms, kisfrekvenciás oszcillátorának sebessége 1-5 Hz. Ez olyan sok és ráadásul relatív gyorsan változó helyű beszakadást okoz a frekvencia-átvitelben, hogy ott már nem azt érezzük, hogy a hangszín megváltozott, hanem csak gazdagabbnak, teltebbnek halljuk a hanganyagot. Ezzel szemben a flanger jól széthúzott beszakadási frekvenciái (kb. 1-5 ms-os átlagos késleltetési időhöz tartozó beszakadások) jól hallható hangszínmódosulást eredményeznek, ráadásul a lassan változó (0.3-1 Hz) késleltetés-módosulást pedig már nem érezzük hangmagasság-változásnak.

A mai flanger-ek általában visszacsatolást is tartalmaznak, mely tovább növeli a hatást (ld. 2.20. ábra).

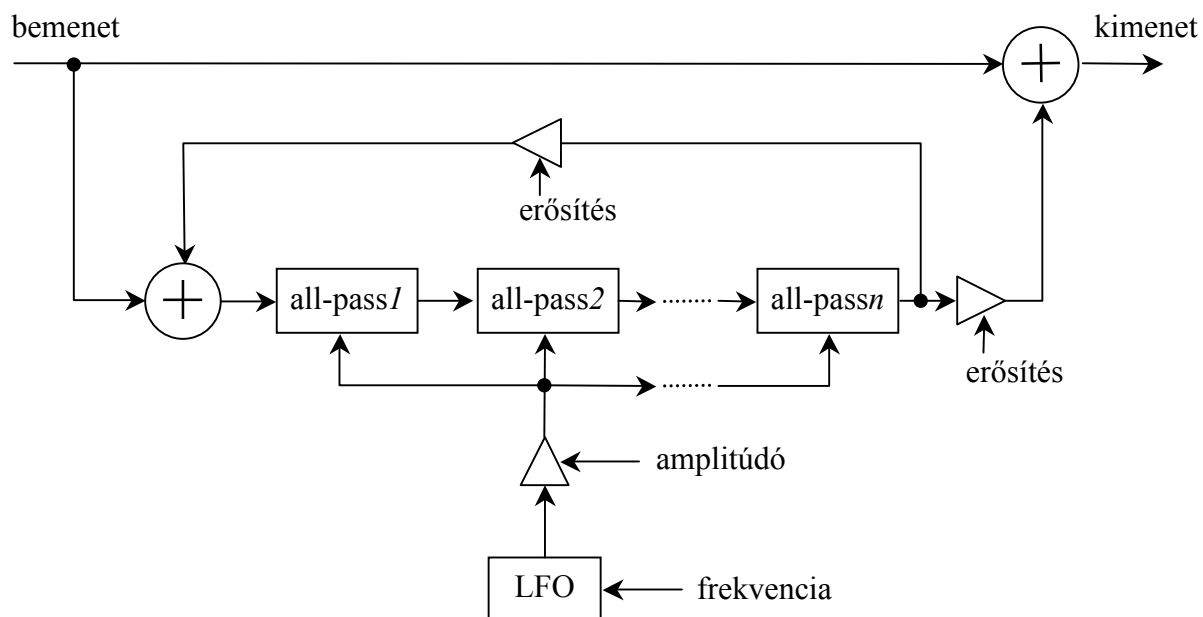


**2.20. ábra. A flanger sematikus rajza**

Az alacsonyfrekvenciás oszcillátor (LFO) frekvenciájának megadásával, majd az oszcillátor által kiadott (pl. szinuszos) jel amplitúdójának beállításával kaphatunk egy olyan vezérlőjelet, mely az elhangolás sebességét (LFO frekvencia) és mélységét (LFO jelének amplitúdója) egyaránt vezérli. A DC eltolás az átlagos késleltetési időt adja meg. Vegyük észre, hogy ez pontosan ugyan az, mint a kórusnál alkalmazott késleltető + elhangoló elemek. A különbség csupán annyi, hogy itt az effekt jellege miatt nem elhangolásnak nevezik a változó sebességű lejátszást.

## Fázistoló – phaser

Ha az eredeti jelhez annak pusztán fázisában módosított változatát adjuk hozzá, akkor kapjuk a fázistoló vagy közismertebben phaser effektet (ld. 2.21. ábra). Ennek megvalósítására mindentáteresztő (all-pass) szűrőket használnak. Hasonlóan a flanger-hez, egy alacsonyfrekvenciás oszcillátorral a fázistolókban is „mozgatják” a beszakadások középfrekvenciáját, valamint általában visszacsatolást is tartalmaznak a kapcsolások.



2.21. ábra. A fázistoló sematikus rajza

## Vibrató – vibrato

A hegedűn való játékmódot utánzó effekt, amely az eredeti jel helyett olyat ad ki, amely periodikus elhangoláson ment keresztül. Ez nagyon hasonlít a kórusban alkalmazott megoldáshoz (ld. 2.18. ábra) azzal a különbséggel, hogy itt a direkt jel nem kerül rá a kimenetre, és így módon a késleltetőre sincs szükség. Paraméterek tekintetében elmondható, hogy itt valamivel nagyobb frekvenciájú a hangmagasság-változtatás.

## Tremoló – tremolo

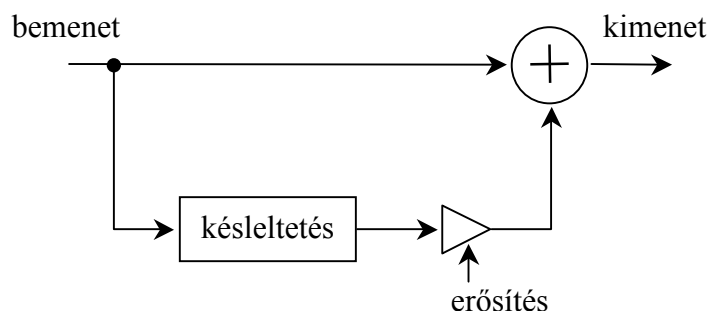
A tremoló talán az egyik legegyszerűbben megvalósítható effekt: az előzőekhez hasonlítva a kisfrekvenciás generátor itt nem egy hangmagasság-eltolót, hanem egy hangerőszabályzót vezérel. Hatása a remegő hanghoz hasonlítható leginkább.

## Sztereoó hatás

Képzeld el, hogy egyszál monó jelből szeretnénk széles sztereoóhatású hangzást szimulálni. A fülünk akkor hinné ezt, ha a két fül mást-mást hallana, de mégis úgy éreznénk, hogy középről szól a hang. Ezt úgy lehet elérni, hogy az eredeti jelet az egyik csatornára vezetjük, a másikat pedig 10-20 ms



késleltetéssel a másik csatornára. Mivel fülünk a korábban meghallott hangra lokalizálna és egész egyszerűen csak úgy éreznénk, hogy oldalról jön a hang, azért addig erősítjük a késleltetett jelet, amíg újra középről nem halljuk. Ha ui. ugyanabban a pillanatban ugyanaz a jel egyik fülünkbe hangosabban jut be, akkor arra az oldalra lokalizálunk.

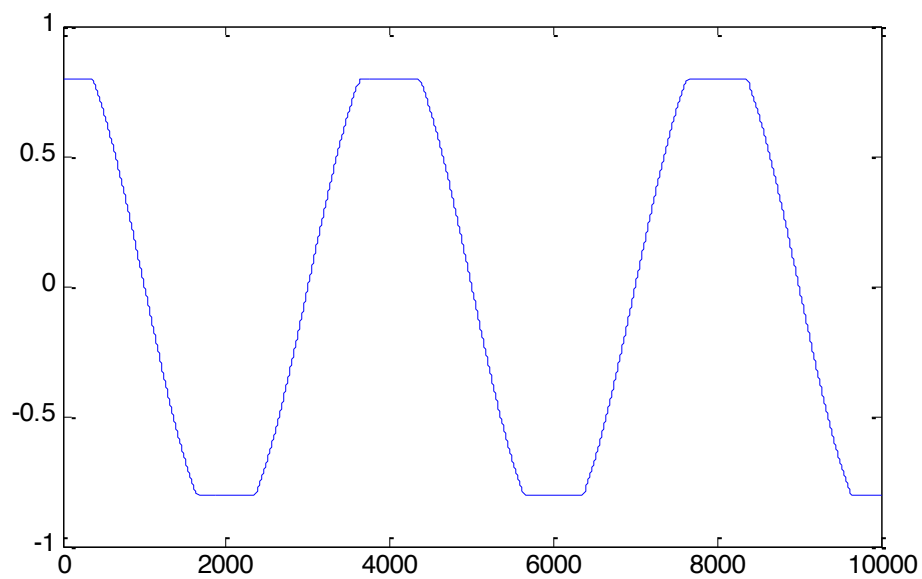


**2.22. ábra. Sztereó hatású effekt funkcionális rajza**

A késleltető ideje 10-20 ms, ez erősítés 6 dB körüli

## Torzító

Gitárosok körében kedvelt effekthez jutottunk. A lényeg az, hogy a jelgörbe tetejét valamiképpen megtörjük, és ezáltal különleges hangzást kapjunk. A legegyszerűbb megvalósítás digitálisan egy egyszerű jelszint-korlátozással érhető el (ld. 2.23. ábra), hagyományosan, analóg technikával egy jelentős mértékben túlvezérelt műveleti erősítőt szoktak alkalmazni. Innen a neve az Overdrive gitárpedálnak. A „megtörés” lehet jelszint-független – ez az egyszerű túlvezérlés. Lehet jelszint függő is (ezt az effektet Fuzz-boksz néven szokták emlegetni), ezzel elérhető, hogy a gitáros hangosan és halkán is játsszon egyaránt, a jel mindig ugyanolyan mértékben lesz torzított. Aztán lehet az egyszerű vágás valamilyen egyéb kicsit simítottabb formában végrehajtani. Pl. egy soft-knee-vel beállított limiterrel (ezt analóg technikában a visszacsatoló ágba helyezett diódákkal érik el). A megvalósítások tárháza kimeríthetetlen, mindenki másra esküszik. A „menő” az, ha a csöves hangzást valljuk...



2.23. ábra. Túlvezérelt jel

### 3. Keverés

Az előző fejezet arra koncentrált, hogy egy felvételi csatornán milyen jellegű módosításokat lehet végrehajtani. Ha a kész felvétel több csatornából tevődik össze egyetlen (sztereó, többcsatornás) felvétellé, akkor meggondolandó, hogy bizonyos effektek hatása a többi csatornával való együttes lejátszás során állítandó be. Tipikusan elektromos hibák és vágási munkálatok önállóan is végrehajthatóak, mások azonban, főként a hangzást befolyásoló beavatkozások, csak a teljes összhangzáshoz viszonyítva végzendők.

A számítástechnika rendkívül gyors fejlődésével mára már gond nélkül képesek lehetünk otthoni számítógépünkkel 10-16 csatornát (2002-es adat. 2007-ben min. 40-50 csatornát tudunk) egyidejűleg lejátszani és effektek tömkelegével ellátni (kivétel, ha konvolúciós zengetőket használunk...). A virtuális stúdiótechnika is nagyon magas színvonalra jutott, mára már szinte minden céleszközt megvalósítottak szoftveresen is. Abba a vitába nem szeretnék belemenni, hogy rosszabbul szól-e, mint az igazi berendezésekkel készült felvétel – technikailag nézve nem sok különbség van a számítógépes keverés és mondjuk egy digitális keverőpult között.

A konkrét témák előtt mindenképpen szükséges néhány fogalom definiálása ahhoz, hogy egyértelmű legyen, mikor miről van szó.

- (hang)állomány: winchesteren tárolt hanganyag, hangminták egymásutánja. Fontosabb paraméterei: a mintavételi frekvencia, a számábrázolási pontosság (felbontás, pl. 16 bites), csatornák száma: monó (1), sztereó (2), több.
- sáv: az a logikai egység, ami egyetlen állományt befogad.
- csatorna: felvételi vagy lejátszási útvonal; egyetlen monó hanganyag lejátszását vagy felvételét teszi lehetővé **vagy** felvételi/lejátszási egység: monó hanganyag.

Ennek megfelelően

- egy hangállomány több csatornát is tartalmazhat

- egy sáv lejátszása több csatornát is igénybe vehet a virtuális világban
- felvétel során több fizikai bemenet ugyanannyi csatornán jut be a számítógépbe, de ezek hangállományonkénti felosztása változhat. Pl. 2 sztereó bemenettel rendelkező kártya jele egyik sztereó bemenete egyetlen sztereó hangállományba tárolódik, míg a másik sztereó bemenet csatornáin érkező jelek külön-külön monó hangállományokba kerülnek. Ennek eredményeképpen felvételkor egyszerre 3 hangállomány íródik: egy sztereó és kettő monó.
- lejátszás során a virtuális csatornák tetszőlegesen keverhetők egymásba és tetszőlegesen konfigurálható, hogy mi jelenik meg a fizikai kimeneteken: egyes csatornák jelei, csoportjelek, avagy a teljes keverés.

### 3.1. Előkészületek

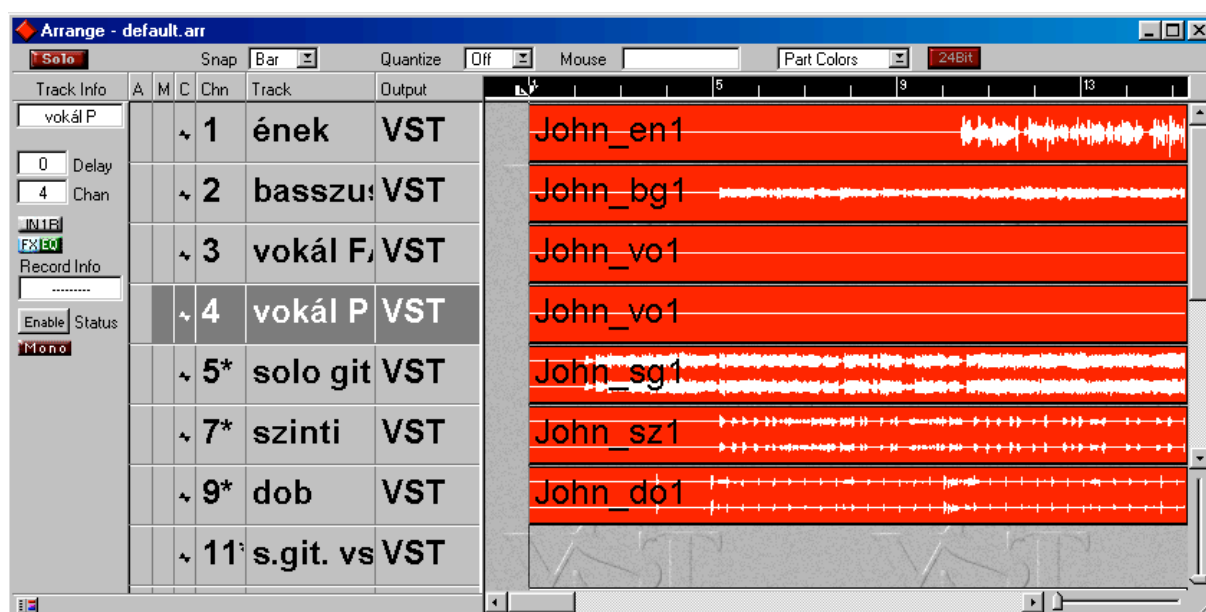
Többcsatornás felvételnél a legfontosabb dolog, hogy a lejátszandó sávok egymáshoz képest szinkronban legyenek. Amennyiben a sávokat külön-külön vesszük fel, úgy általában az előadók segítése érdekében a rájátszásos módszert alkalmazzuk. Ez azt jelenti, hogy a már meglévő sávokat – még ha a legalapvetőbb szinten, effektek nélkül, de – összekeverve lejátszuk az előadónak, aki pedig a zenét hallva arra rájátszik. Technikai oldalról nézve itt a legfontosabb követelmény a szinkronban történő felvétel és lejátszás. Vagyis, hogy a hangállományokban ki lehessen jelölni egy olyan pontot (tipikusan az elejét), ami egy időpillanatban történt a többi hangállomány megjelölt pontjaival. Ez lehet például beszámolás, vagy taktjel az állomány elején, amit egyszer felveszünk, a többi állományba pedig felvételkor mindig belekeverjük a lejátszó eszközről, majd röviddel a zene megszólalása előtt már némítjuk. Ez elegendő lehet arra, hogy minden állomány elején megtaláljuk ugyanazt a beütést. Ha megvan, akkor a hosszabb előfelvételi szakasszal rendelkező állományok elejét kivágjuk úgy, hogy a beütés előtt minden állomány ugyanannyi hangmintát tartalmazzon. Bár sok többcsatornás program támogatja a szinkron felvételt, de ez többnyire csak azt jelenti, hogy a lejátszás megkezdése előtt valamivel

korábban elindítja a felvétel. Majd valamilyen beállítható/automatikus korrekciós idő után felhelyezi a felvételt a kívánt sáv megfelelő pozíciójába többé-kevésbé pontosan. A gond csak akkor van, ha az egyszer elmozdul onnan. Mivel fogalmunk sincs mennyivel korábban indult a felvétel, ezért rendkívül nehéz és igen pontatlanul tudjuk csak „visszatoligálni” a hangállományokat a pontos helyükre.

Digitális felvételnél külön szempont, hogy az egyes csatornákat ugyanazzal az órajellel rögzítsük, hiszen a keverés után az összegjel egyes komponensei (a csatornák) mind egyugyanazon órajellel lesznek lejátszva. Hogy ez mennyire fontos, azt a következő példával szemléltetném. Vegyünk alapul egy 3 perces felvételt, ez 44100 Hz-es abszolút pontos mintavételi frekvenciával. Ha ezt most egy 100 ppm-es, azaz 0,0100 % hibájú, lassabb mintavételi frekvencián lejátsszuk, úgy a lejátszási mintavételi frekvencia 44096 Hz lesz (a 44100 Hz 99,99 %-a). Ez a mintavételi időközök és így a teljes felvétel hosszának 1/99,99 %-os (0,01 %-os) növekedéséhez vezet; a lejátszási idő 18 ms-al lesz hosszabb, mint amennyinek kellene lennie. Így készítve a felvétel úgy fogjuk hallani, hogy „szétesett” a zene, mivel az 50 ms-on túli időkülönbséggel beérkező hangokat (különböző felvételi csatornákból) már elkülönülten halljuk. Ezért nem lehet több csatornát egyidejűleg felvenni pl. két hangkártyával, hacsak nem rendelkeznek a kártyák külső órajel-szinkronizálási lehetőséggel (ez azonban nem jellemző a hagyományos hangkártyákra). Ezért kénytelenek vagyunk vagy többcsatornás hangkártyát vásárolni, vagy a mindenképpen több csatornán rögzítendő hangszereket (pl. dob) stúdióban, esetleg kölcsön felszereléssel felvenni.

Amennyiben a felvett hangállományok egymáshoz képest szinkronban vannak, és ugyanabban az időpillanatban kezdődnek, úgy nem marad más dolgunk, mint a virtuális stúdió program lejátszási sávjaihoz „beemelni” őket (3.1. ábra). Fejlettebb programok számára a sávok tartalma sokféle lehet: monó, sztereó, avagy többsávós hangfelvétel, midi állomány. A formátumot a program felismeri, majd lejátszáskor a megfelelő szoftvert alkalmazza hangkibocsátáshoz. Ez kicsit más szemlélet, mint amit a stúdiókban lehet látni, ahol hang esetén 1 sáv minden esetben 1 csatornát jelent, itt 1 sávra egyetlen összetartozó egység

kerülhet. Visszatérve a fejezet elején bevezetett fogalmakra, a sávokról elmondható, hogy azok mindig szinkronban vannak, azok csak „vízszintes csíkok”, melyekre hangállományokat lehet „felpakolni”. A hangállományok a sávokon belül viszont szabadon pozícionálhatók – szabadon csúsztathatók ide-oda. Amint az a 3.1. ábrán látható ez az elrendezés 4 darab mónó hangállományt tartalmaz 4 sávon – egyenként 1-1 csatornát elfoglalva a keverőpulton, és tartalmaz 3 sztereó (2 csatornás) hangállományt 3 sávon – melyek egyenként 2 csatornát foglalnak el a keverőpulton.



3.1. ábra. Virtuális stúdió program által lejátszandó sávok

A sávok tartalma szabadon mozgatható, de könnyebb dolgunk van és sokkal pontosabb a szinkron, ha a hangállományok kezdete egymással szinkronban van. Ilyenkor minden állomány elejét egyetlen ponthoz kell igazítani.

### 3.2. Lejátszás indul – PC-vel!

A PC-s hangállomány tárolás olykor megfelelő figyelmet követel. Lejátszás során ui. a lejátszó program úgy működik, hogy mindig egy adott szeletkét beolvas először az első hangállományból, majd a másodikból, harmadikból stb. és mikor minden a memóriában van, akkor elkezd a feldolgozást, keverést. A szükséges feltétel ahhoz, hogy a dolog működjön, az első lépésben az, hogy a winchesterről elegendően gyorsan megérkezzenek az adatok. Hogy ez megtörténik-e, a meghajtó sebességparaméterein és az állományok töredezettségi állapotán egyaránt múlik. A meghajtó paramétereinek közül a legfontosabb az

olvasási sebesség, amit többnyire folyamatos, egybefüggő állományok olvasására értenek. Még az sem mindegy, hogy az állomány a merevlemez belső, vagy külső területén helyezkedik el (fizikailag). Ezért többnyire megadnak átlagos olvasási sebességet. Figyelni kell arra, hogy a winchesterek először a gyorsabb elérési területekre írnak, majd amikor kezd megtelni a meghajtó, akkor jut a lassabb részekhez. Több partíció esetén az egymás után létrehozott partíciók a létrehozás sorrendjében egyre lassabban érhetőek el. Következő szempont a winchester elérési idő adatai: az egybefüggő olvasást csak a sávról-sávra (track to track) ugrás lassítja, viszont új állomány, új szakasz megkeresése teljes fejpozicionálást igényel, ami lényegesen lassabb. Ezért ha az állományok össze-vissza feldarabolódva kerültek fel a winchesterre, akkor a folyamatos kiolvasás lassú lesz, hiszen rendszeresen kell újrapozicionálni az olvasó fejet. Ilyenkor segíthet a töredezettség-mentesítés. Ha minden rendben, a winchester gyors, az állományok relatív egybefüggők, akkor már csak a busz átviteli sebessége lehet a korlát. A korábbi PIO (Programmed Input Output – programozott be és ki olvasás/írás) üzemmódhoz képest jelentős előrelépést hozott az UDMA (Ultra Direct Memory Access – ultra (gyors) közvetlen memória elérés) a maga 33 MB/s-os átviteli sebességével. Ezt ezután tovább gyorsították és megjelentek az UDMA 66-os, 100-as és mára már 133-as adatátvitelt támogató meghajtók és alaplapok. Ez az átviteli buszsebesség-különbség szintén megvan a professzionális – SCSI – meghajtóknál is (SCSI, Wide SCSI, Ultra Wide SCSI).

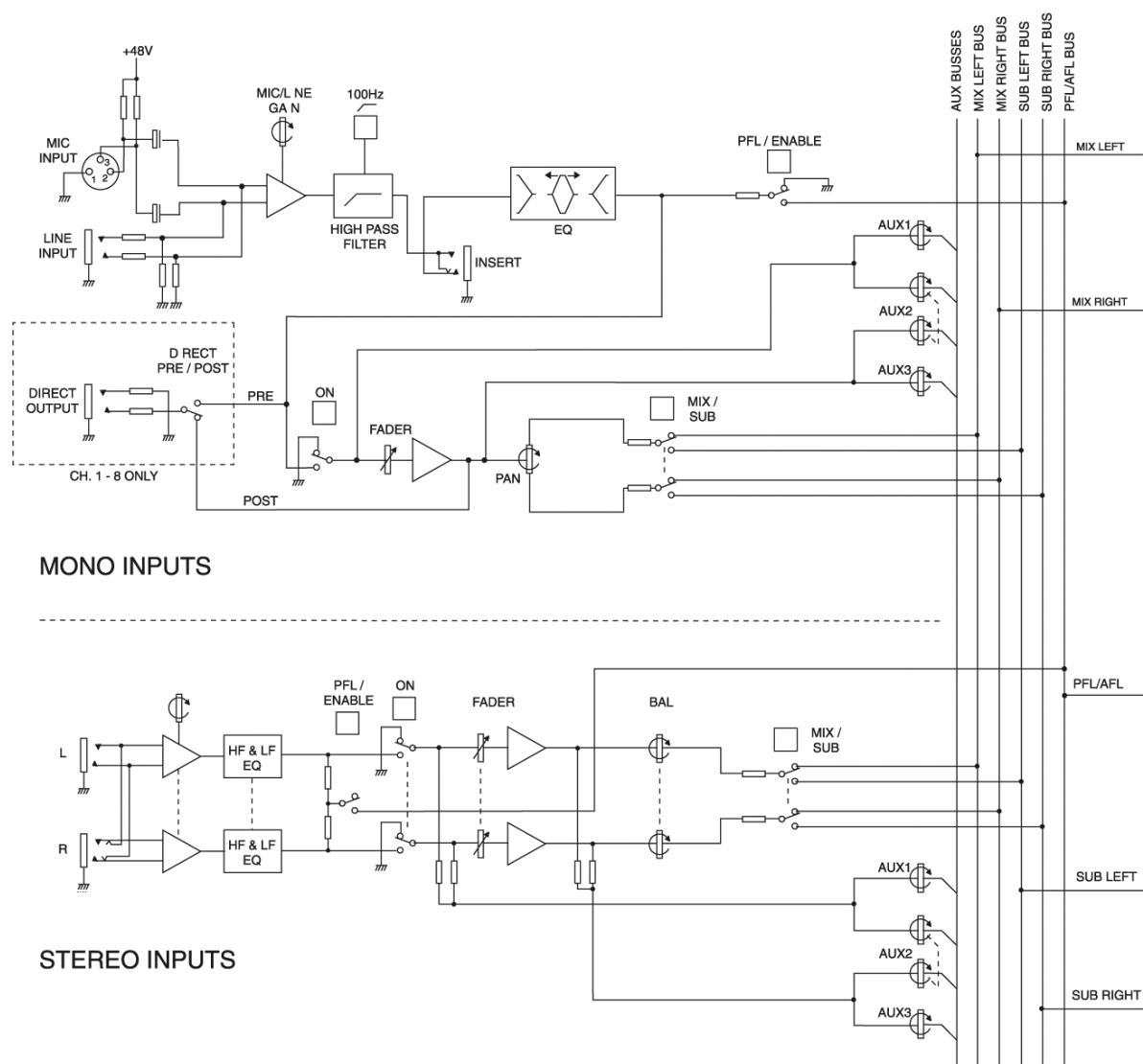
### 3.3. Keverés

A következőkben egy virtuális keverőpult elé fogunk leülni és áttekintjük hogyan is lesz az egyenkénti csatornákból kész, kiadható felvétel. A mostani tárgyalásmód szempontjából szinte érdektelen, hogy igazi keverőpult előtt ülünk vagy a számítógépes virtuális keverőpult előtt, kezelni lényegében ugyanúgy kell. Sokszor a virtuális keverő szemléletesebb, mert mindent grafikus felülettel képes ellátni, ami sok hardveres megvalósításnál (pl. effekt berendezések) – még a drágább kategóriákban sem – található meg.

Ebben a fejezetben elsősorban az egyszerűbb tárgyalásmód érdekében egyetlen virtuális stúdió szoftveren keresztül fogom bemutatni a keverés alapjait,

de ez semmiképpen sem akar reklám lenni, hanem számomra ezen a programon a legkényelmesebb bemutatni. Ez nem jelenti azt, hogy ez a legjobb program, mindenkinek javaslom, hogy mielőtt vásárol, próbáljon meg kipróbálásra többfélét is megszerezni, és csak azután kötelezze el magát.

A keverőpultokról a legtöbbet a jelutakat ábrázoló leírásuk árul el, és a számítógépes feldolgozás is lényegében ugyanígy megy végbe. Ezért nézzük is mit lehet leolvasni egy ilyen ábráról:



3.2. ábra. Analóg keverő jelút-diagrammja: csatorna szekció

## Analóg keverő jelútjai

Az első észrevevő az, hogy a sztereó csatornák csak annyiban különböznek a monó csatornáktól, hogy a bal és a jobb oldali szabályzók össze



vannak követve. Ez még természetesebb a virtuális technikában, hiszen ott semmiből sem áll olyan csatorna csoportokat – nem feltétlenül csak kettő csatornát – létrehozni, amelyek közös vezérlést kapnak. A következőkben tehát alapvetően mindig egy csatornáról fogok beszélni, de a műveletek értelemszerűen több csatornára is vonatkozhatnak.

A jelbemenet először is kétfajta lehet, mikrofon, avagy vonali szintű. A megfelelő bemenet alkalmazása után a jel egy előerősítő fokozatra kerül, mely arra hivatott, hogy a kimenetén előálló szint az összes csatornán kb. azonos legyen. Ez a „szintezés” arra jó, hogyha a csatorna-hangerőszabályozóját két csatornának azonosra állítjuk, akkor azok kb. azonosan hangosak is legyenek. Ezután a jel egy 100 Hz-es felüláteresztő szűrőn megy keresztül, ha aktiváljuk és ha a keverő rendelkezik ezzel az opcióval. Ennek célja az énekbemenetek káros kisfrekvenciás zajainak eltávolítása (szél, lépés stb.). Ez az a szakasz ameddig a jel mindenképpen analóg módon megy. Az első fokozat különösen kritikus, mert az ott felszedett zaj és egyéb torzítások többet nem kerülnek ki a rendszerből. Ezután következik az a pont, ahonnan kezdve virtuális stúdiók kezdődnek. Digitális keverőpult esetén itt vagy az ezt követő insert pont után mindenképpen jön az analóg/digitál átalakító.

A következő fokozat az insert pont. Az átviteli láncba itt illeszthetünk be olyan effekteket, amelyek az adott csatornára egyediek. Pl. egy kompresszor, esetleg egy torzító, egy felharmonikus dúsító, speciális szűrők stb. Mindenképpen olyan effekt, mely nem közösíthető a többi csatornával. Pl. az egész keverés térhatását megadó zengető tipikusan először összegzi az összes csatorna zengetésre szánt jeleit, és az összegjelet zengeti meg. Fölöslegesen növekedne a számítási komplexitás, ha minden egyes csatornát egyenként zengetnénk az insert ponton keresztül – bár lehetne, és külön nehézséget okozna, hogyha a zengető paraméterein állítani akarnánk, akkor azt minden egyes csatorna zengetőjén meg kellene tenni az egységes hangzáskép megtartásának érdekében.

Az insert pont után következik a hangszínszabályozó szekció. Itt a már korábban megismert hangszín-szabályozási lehetőségekkel találkozhatunk.

Hagyományos keverőpultok esetén a komolyabbak parametrikus szabályzókkal rendelkeznek, az egyszerűbbek csak sima magas és mély szabályzókkal.

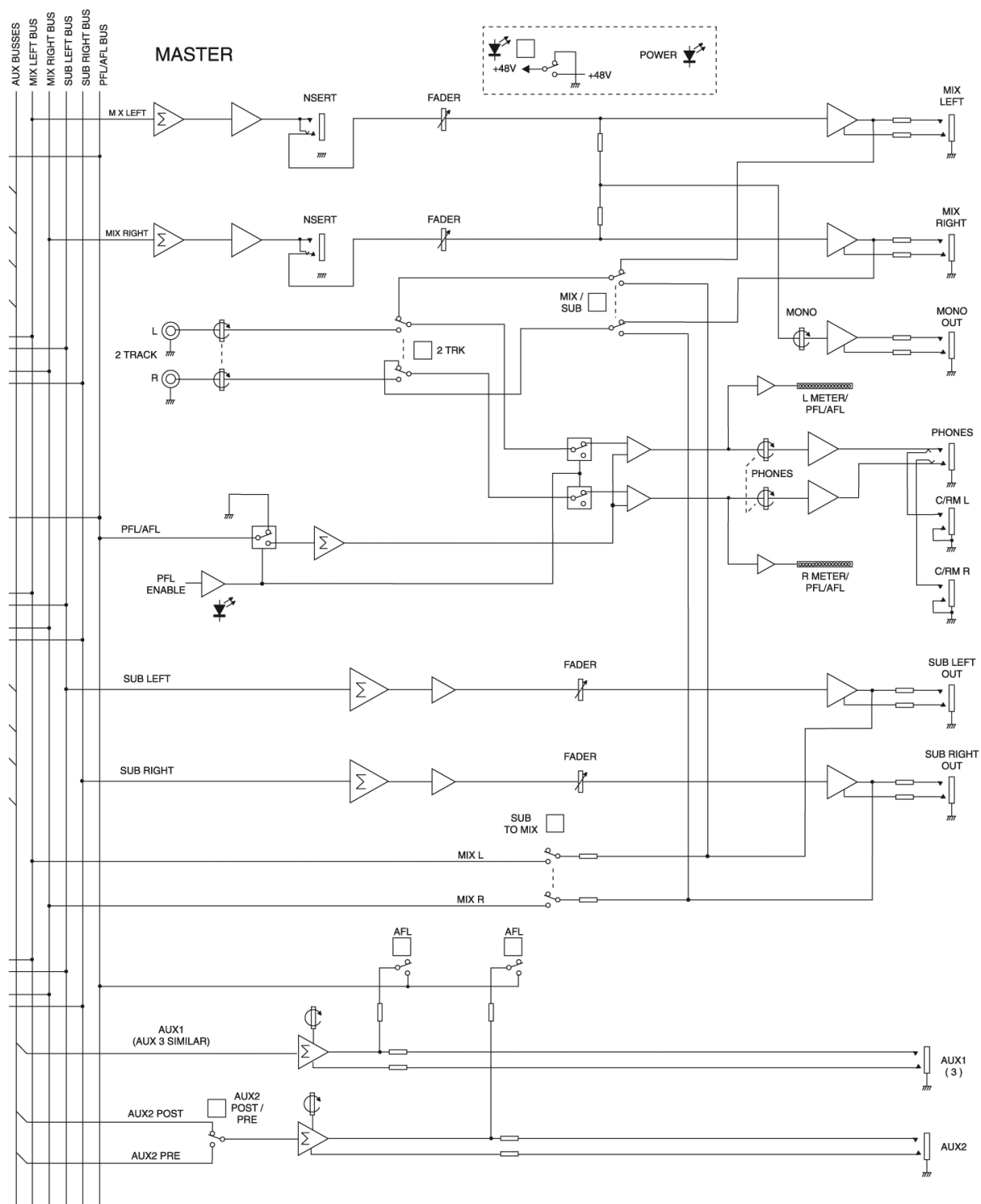
Ha ezen a szekción is áthaladt a jel, akkor már kész a – többnyire normalizált szintű – csatorna-jel. A normalizáláshoz jó segítséget nyújt a PFL (pre fader listening – hangerőszabályzó előtti belehallgatás) felirattal jelölt gomb. Ennek megnyomásával a hanglánc ezen pontjáról – tehát még a hangerő- és panorámaszabályozás előtt – leágaztathatjuk a jelet egy külön csatornára, melyet kifejezetten arra találtak ki, hogy az egyes csatornába bele lehessen hallgatni. Amennyiben a keverőpultunk csak egyetlen kivezérlésjelzővel és fejhallgató kimenettel rendelkezik, úgy ezt a „külön csatornát” kell kiválasztanunk a kivezérlésjelző bemeneteként ahhoz, hogy a csatornákat színtezhessük.

Folytatatva a hangjel útját, ezután következik a térben elhelyezés, majd a keverés. A csatornán tehát beállítjuk a végleges hangerőt, majd elhelyezzük a hangtérben. Sztereó keverés esetén ez bal és jobb csatornába történő szétosztást jelent, surround keverés esetén többnyire 5 csatorna hangterébe helyezzük bele a feldolgozandó csatornát. A hangtérben elhelyezett sztereó, avagy surround jel ezután megválasztható, hogy mely kimeneti pontokra kerüljön rá. Egy master kimenet mindig van, de sokszor találunk több buszkimenetet, amelyekre szabadon irányíthatjuk a jelet. A buszkimenetek elsődleges feladata, hogy a jeleket fogadó eszközök (felvevő magnó/harddiszkrekorder, lehallgató helyiség, feljátszó helyiség) különböző alcsoportjait kaphassák meg a kevert jeleknek.

Nem keverendő az AUX megjelöléssel illetett segédlánc-kimenetek az imént említett buszkimenetekkel. Ezek ugyanis egészen más célt szolgálnak. A leágazási pontjuk két helyen lehet, a csatorna-hangerőszabályzó előtt vagy után. Az angol terminológia: pre ill. post fader. A csatorna-hangerőszabályzó előtti leágazás arra jó, hogy a fő keverés mellett egy másik keverést is megvalósíthassunk. Pl. a színpad elején álló énekes egészen mást kell, hogy halljon, mint amit a közönségnek szánunk kifelé. Az énekes közelében ott szólnak a környező hangszerek, melyek lokálisan elég hangosak, kifelé azonban nem. Ráadásul gyakran megvan, hogy az énekes avagy az adott hangszer számára mely szólamok a legfontosabbak. Pl. a dob és a basszusgitár különösen

jól kell, hogy hallják egymást, mert kettejük összehangolt ritmizálása adja meg a teljes zene alapját („rossz alapozású ház pedig összeomlik”). Az énekes számára viszont a telt hangzást megadó (szintetizátor, vokál, ritmus gitár) hangszerek a fontosak, hogy mindig tisztán tudjon énekelni. A csatorna-hangerőszabályozó *előtti* leágazás tehát különböző monitorozási keverések megvalósítására való, mivel az ezen a kimeneten megjelenő jel független a csatorna fő hangerőszabályozójától. A csatorna-hangerőszabályozó *utáni* leágazás ezzel szemben olyan effektek használatát támogatja, melyeket egyszerre több csatorna is igénybe vesz. Ilyen pl. a zengető. Ha a zengetőt úgy állítjuk be, hogy csak a zengetett jelet adja ki, akkor a post-fader segédlánc kimenettel meghatározhatjuk a közvetlen és a visszavert (azaz a zengető kimenete) jelek arányát. Ez szintén egyfajta térbeli elhelyezésre ad módot. Viszont a fő csatorna hangerő-szabályozó változtatásával a visszavert és a direkt jel aránya nem változik, a hangszert tehát nem érezzük úgy, hogy előrébb vagy hátrébb ment volna.

A különböző összegzett jelek ezután újra átmehetnek a teljes hangfeldolgozási lánc minden elemén (insert pont, hangszín, stb.). A keverőpultoknak ezt a részét nevezik master szekciónak. Itt már csak összegjeleket dolgozunk fel, de a kezelőfelület sokszor teljesen ugyan olyan, mintha közös csatorna jelet dolgoznánk fel.



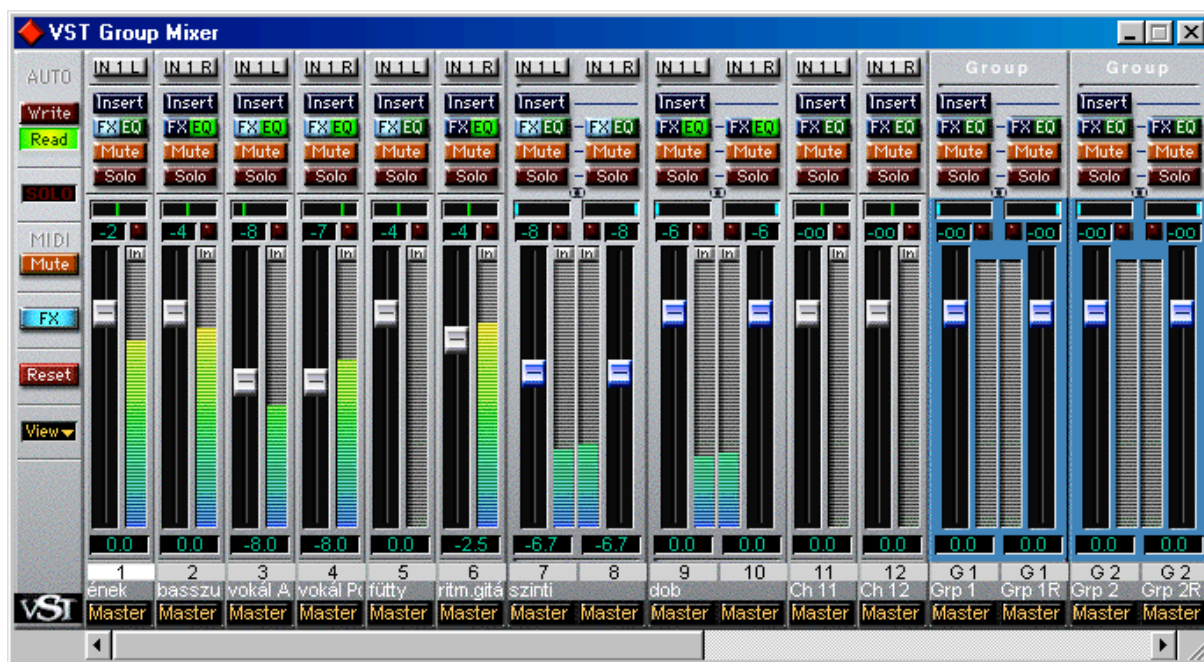
3.3. ábra. Analóg keverő jelút-diagrammja: master szekció

## Virtuális keverők

Most ugorjunk át a virtuális világba, ahol is a fent említett lehetőségek mind rendelkezésünkre állnak, csak sokszor nem annyira egyértelmű – megfelelő

ismeretek nélkül -, hogy mi mit jelent. Viszont hamar rá lehet jönni, ha az előbb hangút leírást megértettük.

A 3.4. ábrán látható virtuális keverő könnyen áttekinthető. Az egyes csatornák bemenetére a jelen rendszerben rendelkezésre álló fizikai 2 bemenet (1 sztereó) valamelyike „köthető” úgy, hogy az adott csatorna bemenetén aktiváljuk az „IN 1 L” ill. „IN 1 R” gombokat. Amennyiben többcsatornás hangkártyával rendelkezünk, úgy természetesen az egyes csatornák bemenetére több fizikai csatorna közül választhatunk, így a virtuális keverőnk igazi digitális keverőpultnak is használható. A csatornák alján található „Master” felirat azt jelöli, hogy feldolgozás után az adott csatorna jele hova megy. Ide megadhatnánk csoport bemeneteket is, de megadhatunk fizikai kimeneteket is.



3.4. ábra. A virtuális stúdió keverő pultja – csatorna és csoport szekció

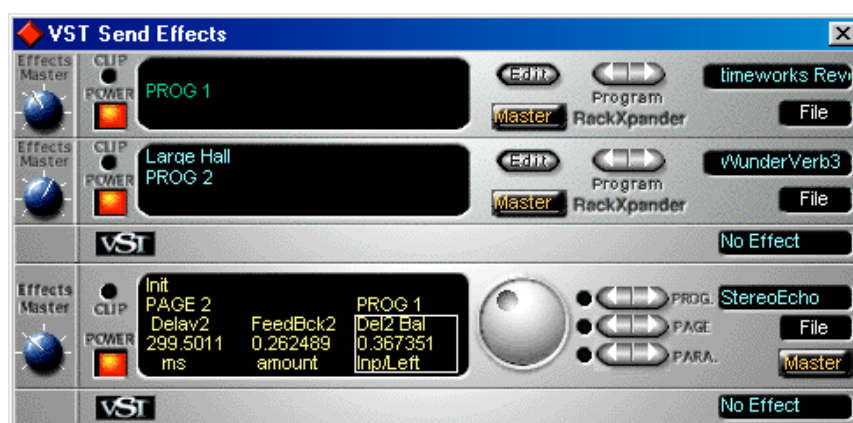
Visszaugorva a csatorna bemenetére láthatjuk, hogy jön az „Insert” pont, majd az effektek, a hangszínszabályozás, a panoráma- és a fő hangerő szabályozás. Az eddig nem említett gombok a „Mute” (némítás), a nevének megfelelően az adott csatorna némítását teszi lehetővé a hangerőszabályozó teljes lehúzása nélkül. A „Solo” gomb arra szolgál, hogyha csak egyetlen egy csatornát szeretnénk a fő kimeneten hallgatni. Ez kicsit hasonlít a PFL gombra, a jelentős különbség azonban az, hogy a PFL gomb a belehallgatni kívánt csatornák jelét egy külön kimeneten gyűjti, míg a Solo gomb lényegében csak azt

jelenti, hogy az összes többi csatornán lenyomjuk a némítás gombot. Bár az analóg keverőpult leírása során nem esett róla szó, a Solo és a Mute gombok komolyabb analóg keverőpultokon is megjelölhetők.

Az Insert pontra rákattintva egy folyamatosan bővíthető effekt-rackot láthatunk, melyeket telepakolhatunk igényeinknek megfelelően (ld. 3.5. ábra).



3.5. ábra. Egy csatorna Insert pontjára kötött Effekt-rack



3.6. ábra. Csatornaszabályzó utáni segédlánc-kimenetekkel meghajtott („közös”) effektek

A 3.7. és a 3.9. ábrákon a virtuális keverő egyetlen csatornájának segédlánc-kimeneteinek vezérlését ill. a hangszínszabályzó szekciókat láthatjuk. A virtuális világ könnyebb konfigurálhatósága miatt itt többnyire alig van korlátozás, jelen esetben minden csatornán található 8 segédlánc-leágazás, melyekről egyenként eldönthetjük, hogy éljen-e vagy sem (On kijelző), hogy hangerőszabályzó előtti vagy utáni legyen a leágazás (Pre kijelző), hogy mekkora hangerővel és melyik effektekre menjen a leágazás. Az egyes csatornákhöz felépített, Insert pontra kötött effekt-rackon kívül még egy effekt torony építhető,



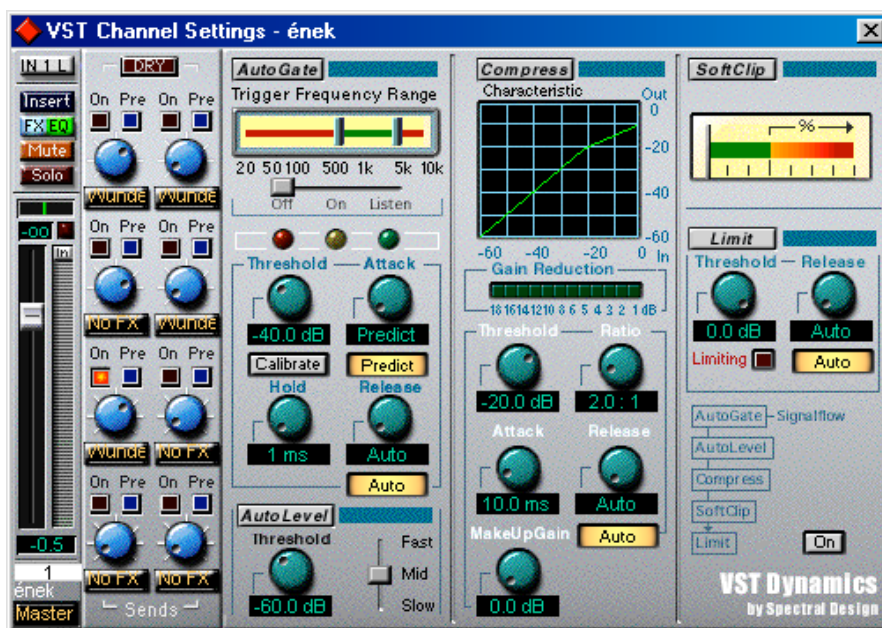
melyeket közösen használ az összes csatorna (ld. 3.6. ábra). Ennek az effektjeit lehet a segédlánc-kimenetek megfelelő beállításával meghajtani.

Hangszínszabályozás tekintetében (3.7. ábra) itt nem lehet panaszunk a konfigurálhatóságra: teljes mértékben paraméterezzhető hangszínszabályozókat találunk. Ám aki ezzel a felülettel, nincs kielégedve, avagy speciális igényei vannak, mint pl. a beállított hangszíngörbe grafikus megjelenése, az szabadon illeszthet be az insert pont utolsó effektjének bármit, a példánál maradva egyedi hangszínszabályzó modulokat (3.10. ábra).

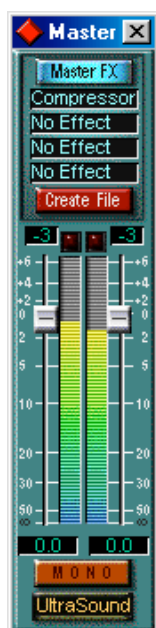


3.7. ábra. Egy csatorna feldolgozó lánc – hangszínszabályozással

A kompresszor és limiter effektek annyira gyakoriak, hogy komolyabb keverőpultokba – a hangszínszabályzó modulhoz hasonlóan – minden csatornába beépítik. A virtuális világban (szoftveresen) könnyen megvalósítható a csatornánkénti kompresszor és a gyakori használatra való tekintettel a virtuális keverő részeként találjuk meg, nem szükséges csatornánként beilleszteni az insert pontokra. Aki persze nem lenne elégedette a „beépített” kompresszorral avagy annak kezelőfelületével, az természetesen használhat speciálisat az insert pontokon.



3.9. ábra. Egy csatorna feldolgozó lánc – dinamikaszabályozással

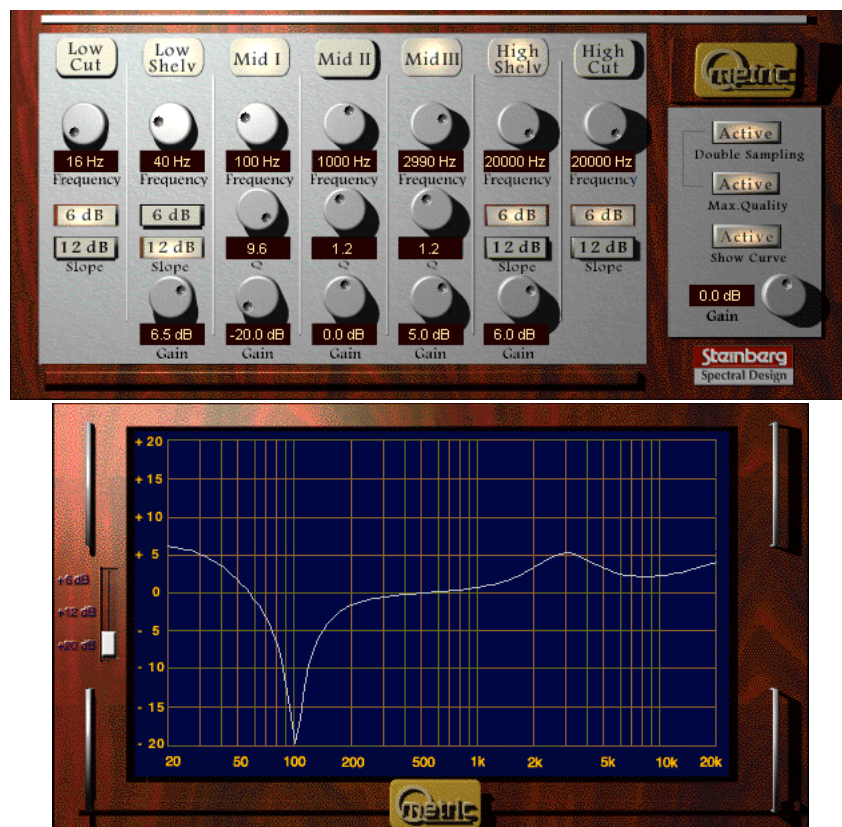


3.8. ábra.  
A master szekció

A master szekció, amint arról már korábban esett szó, lehet ugyanolyan felépítésű, mint a többi csatorna. Ennél a programnál kettéválasztják master szekció: a csoport jeleket gyűjtő csatornákra és a master csatornára. A csoport jelek arra szolgálnak, hogyha pl. a 6-8 csatornán felvett dobfelszerelés egyes komponenseit egymáshoz képest jól beállítottuk, akkor az összes dobos tartalmazó csatornát egyetlen csoport csatornába irányítva (és nem a master-re) megtehetjük, hogy egy hangerőszabályzóval szabályozzuk a teljes dobfelszerelés hangerejét. Más programoknál ezt sokszor úgy oldják meg, hogy ki kell jelölni azon szabályzókat, melyeket közös vezérléssel látunk el, majd ezek közül egyet, amelyik „irányít”. Ha ezt szabályozzuk, akkor „megy vele” az összes többi, ha bármelyik másikat, akkor a csoporton belüli arányok változnak. A csoport-csatornáktól egy kicsit más felépítésű ennek a programnak a master-csatornája. Elsőként azt látjuk, hogy a master-csatornába összegződnek a „közös” effektek, és ezután kerül vihetjük még keresztül a teljes mix-et egy effekt-tornyon.

Ízelítőként az effektek széles tárházából nézzünk néhányat:

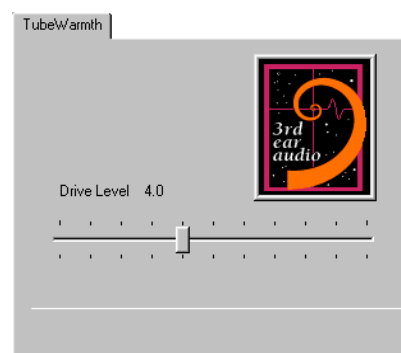




3.10. ábra. Grafikus hangszínszabályozó effekt.



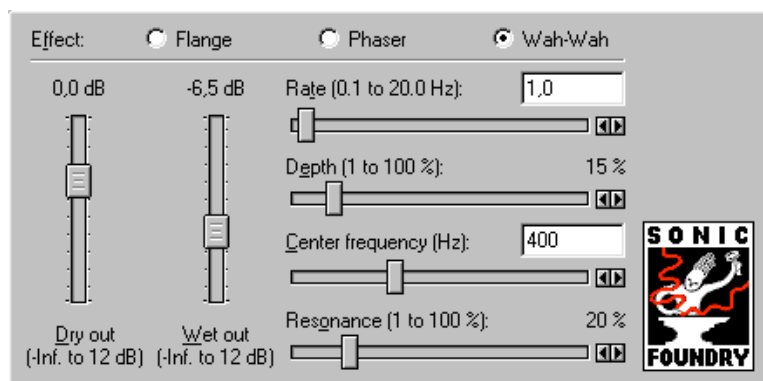
3.11. ábra. Látványos kezelőfelületű zengető



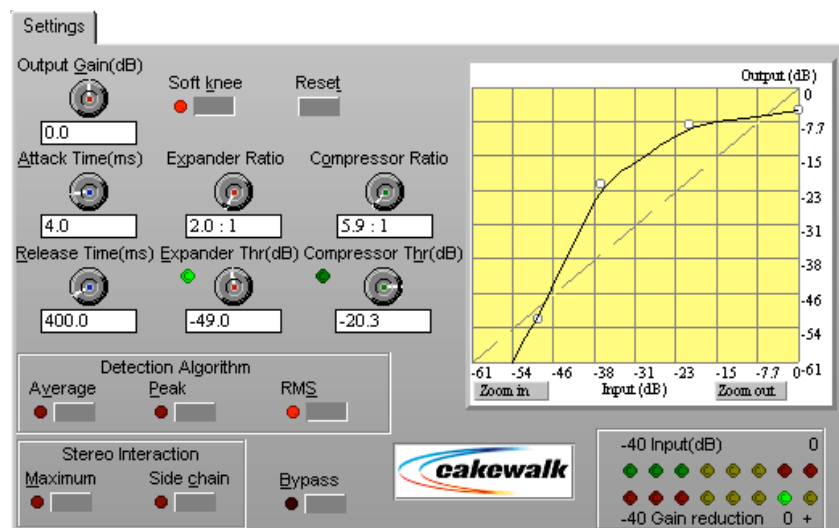
3.12. ábra. Csöves torzító



3.13. ábra. Lemezjátszó zaját utánzó effekt



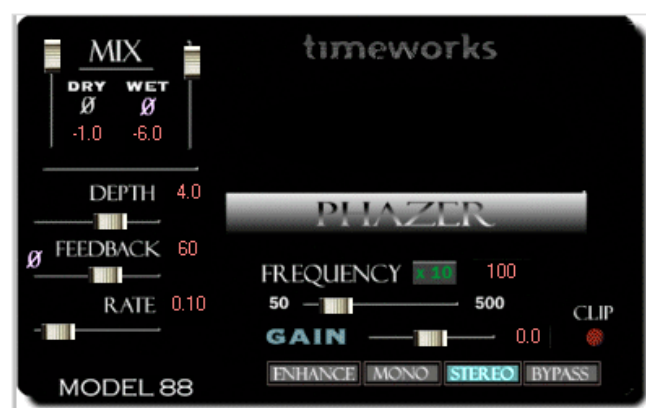
3.14. ábra. Flanger, Phaser és Wah-wah effekt



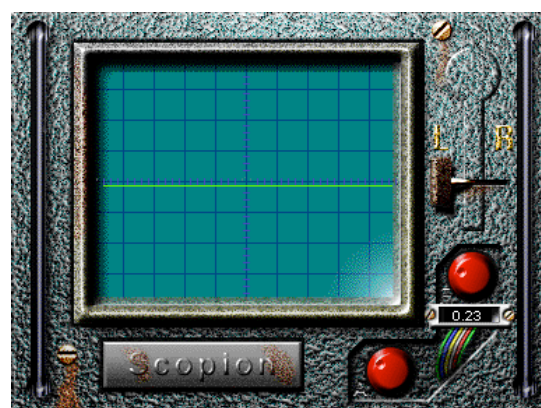
3.15. ábra. Sokoldalú kompresszor-limiter



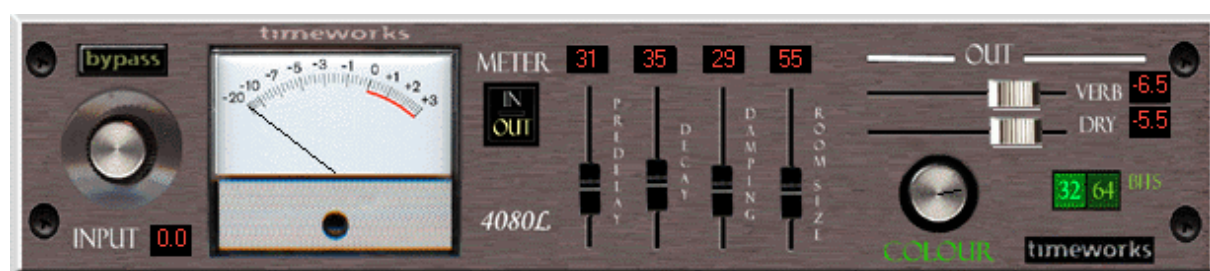
3.16. ábra. Fuzz-box



3.17. ábra. A hagyományörző phaser



3.18. ábra. Egyszerű oszcilloszkóp



3.19. ábra. Az analóg elődjének arculatát őrző nagyon igényes zengető



3.20. ábra. Csak a legalapvetőbb paraméterekkel felruházott zengető



3.21. ábra. Klasszikus effekt rack-be épített kórus

## Automatikus vezérlés

Ha tisztában vagyunk a keverés lehetőségeivel és megkerestük a használni kívánt hangzasképet, akkor már csak finom-módosítások lehetnek hátra, melyeket a lejátszás alatt akarunk véghezvinni. Ilyen lehet például a közönség ovációjának behangosítása a koncertfelvétel egy adott pillanatában, majd fokozatosan visszahalkítani annyira, hogy csak kellemes koncerthangulatot adjon. Digitális keverőpultokon és a szoftveres virtuális stúdió programokban is lehetőség van a szabályzók automatikus vezérlésére. Ez azt jelenti, hogy a vezérlő rendszer képes arra, hogy a lejátszás közben végrehajtott módosítások folyamatát rögzítse, majd a későbbi lejátszások során pontosan úgy végrehajtsa, mint ahogyan mi hajtottuk végre a rögzítéskor. Ugyanolyan ütemezéssel halkít le egy hangerőszabályzót, kapcsol be egy effektet, stb. Később persze felülbírálhatjuk korábbi cselekedeteinket, sőt magukat az eseményeket külön sávon rögzíti a program, melyet azután szövegesen is szerkeszthetünk (3.22. ábra).

| Start-Pos.  | Length | Val.1 | Val.2 | Val.3 | Status   | Chn | Comment      |
|-------------|--------|-------|-------|-------|----------|-----|--------------|
| 0034.04.224 | =====  | 14    | 1     | 13    | AUDIOMIX | ==  | Eq2Enab Off6 |
| 0042.04.260 | =====  | 13    | 1     | 32767 | AUDIOMIX | ==  | Pan 6        |
| 0042.04.260 | =====  | 12    | 1     | 15245 | AUDIOMIX | ==  | Volume 6     |
| 0042.04.260 | =====  | 112   | 0     | 15245 | AUDIOMIX | ==  | Volume 5     |
| 0042.04.261 | =====  | 113   | 0     | 0     | AUDIOMIX | ==  | Pan 5        |
| 0059.01.370 | =====  | 13    | 1     | 16383 | AUDIOMIX | ==  | Pan 6        |
| 0059.01.372 | =====  | 14    | 1     | 77    | AUDIOMIX | ==  | Eq2Enab On 6 |
| 0059.01.372 | =====  | 114   | 0     | 77    | AUDIOMIX | ==  | Eq2Enab On 5 |
| 0059.01.372 | =====  | 12    | 1     | 11377 | AUDIOMIX | ==  | Volume 6     |
| 0059.01.372 | =====  | 112   | 0     | 11377 | AUDIOMIX | ==  | Volume 5     |
| 0059.02.028 | =====  | 113   | 0     | 0     | AUDIOMIX | ==  | Pan 5        |
| 0067.02.226 | =====  | 114   | 0     | 13    | AUDIOMIX | ==  | Eq2Enab Off5 |
| 0067.02.227 | =====  | 14    | 1     | 13    | AUDIOMIX | ==  | Eq2Enab Off6 |
| 0075.02.340 | =====  | 14    | 1     | 77    | AUDIOMIX | ==  | Eq2Enab On 6 |
| 0075.02.344 | =====  | 114   | 0     | 77    | AUDIOMIX | ==  | Eq2Enab On 5 |
| 0083.03.074 | =====  | 14    | 1     | 13    | AUDIOMIX | ==  | Eq2Enab Off6 |

3.22. ábra. A lejátszás-vezérlés tételes listája

## 4. Hangtömörítési eljárások

A hanganyag elkészülte után a következő lépés, hogy milyen médiumon adjuk ki. Digitális hanganyagok esetén elsősorban a CD lemez a megszokott, de mára már a DVD Audio lemez is számításba jöhet. A CD lemezen szabvány írja elő, DVD lemez esetében pedig – hacsak más szempont nincs, akkor – az óriási kapacitás (min. 6 CD-nyi anyag ráfér egy DVD Audio lemezre) miatt tömörítetlenül tároljuk a hanganyagot. Gyakran előfordul azonban, hogy ismerőseinknek, barátainknak Interneten, vagy más kisebb kapacitású hordozón szeretnénk megmutatni a kész felvételt. Ilyenkor többnyire nem is a tökéletes hangminőség a lényeg, hanem az, hogy az anyag eljusson a címzetthez. Ezért általában szóba jön a hanganyag tömörítése is. Kérdés az, hogy milyen módszert alkalmazzunk. E fejezetben részletesen szeretném bemutatni a legismertebb két kódolási eljárást az MPEG Audio-t és az AC-3-at, amit ma már többnyire mindenki csak Dolby Digital-nak ismer. Ezen kódolási eljárások ismeretében később megjelenő kódolási eljárások is könnyen megérthetők lesznek, és így nem lesz nehéz mérlegelni, hogy adott tömörítési eljárással mit „rontunk” el a zenében a kisebb tárigény érdekében.

Hagyományos matematikai tömörítési eljárások nagyon rossz hatásfokkal képesek csak hangjelek tömörítésére. A hangjel típusától függően (hangszer, ének, beszéd, teljes zenekar, stb.) a tömörített állomány az eredeti 60-95 %-a. Így ezek az eljárások nem igazán alkalmasak audió jelek bitsebesség-csökkentésére.

A 70-es évek vége felé kezdtek el pszichoakusztikus hangjelenségeken alapuló kódolással foglalkozni. Ezen eljárások lényege, hogy olyan bitsebesség csökkentési módszert alkalmaznak, hogy a jeltorzulás az emberi hallás korlátozott észlelőképessége miatt nem hallható. Érdemes figyelni arra, hogy itt veszteséges tömörítésről van szó. Ez azt jelenti, hogy az eredeti jelből információt veszünk és ezzel kódolási eljárásunk felhasználhatósági körét csökkentjük. Hangstúdiókban, feldolgozás alatt álló jeleket tehát nem célszerű így tárolni, mert több, egymás utáni kódolás - dekódolás után, netán összekeverve több más hangjellel, végül igen rossz minőséghez juthatunk. Ezzel szemben a nagyközönség számára szánt műsorszórásban (televízió, rádió), ahol

általában egyszeri meghallgatásról van szó, rendkívül hasznos lehet, amennyiben nagy tömörítést tudunk elérni anélkül, hogy a hallgató észrevenné a minőségromlás, vagy az őt zavarná.

Pszichoakusztikus elven működő audio kóderek alkalmazásával nagyfokú tömörítés érhető el, miközben a hallható minőségromlás minimális. Első megközelítésben a bemeneti jelet kritikussáv<sup>1</sup>-léptékű frekvenciatengelyen ábrázolják, majd kvantálják. A kvantálás mértékének meghatározásához spektrális elfedési jelenségeket vesznek figyelembe, hogy a kódolt jel – adott bitsebesség mellett – a lehető legjobb minőségű legyen. A jel spektrális ábrázolására általában valamilyen frekvencia-transzformációt vagy részsávok szűrőbankot alkalmaznak.

Amint azt kutatások kimutatták, az emberi hallás észlelőképesége mind időben, mind frekvenciában korlátozott. A pszichoakusztikus tömörítés lényege abban rejlik, hogy a bemeneti digitális jel mintáinak tárolási pontosságát (bitszámát) úgy csökkentjük, hogy az így a jelhez adódó kvantálási zajt hallásunk korlátai miatt ne észleljük. Minél jobb a tömörítési eljárás, minél pontosabb az alkalmazott hallásmodell, annál jobban kezeli a járulékos zaj idő- és frekvenciabeli lefolyását, annál kevésbé lesz hallható a zaj adott tömörítési fok mellett. Pl. időben lassan változó jelek kódolása esetén a zajjel spektrumának a jel által okozott spektrális hallásküszöb-emelkedés vagy az abszolút hallásküszöb alatt kell maradnia. Spektrális alapon történő bitsebesség-csökkentést széles körben alkalmaznak. Teljesen más a helyzet, mikor a jelszint nagymértékű időbeli változásokon megy keresztül: egy halk hangot egy hangos követ, és megfordítva. Ilyenkor olyan kódolás alkalmazása célszerű, mely gondoskodik arról, hogy a zajjel szintjének időbeli lefolyása mindig az időbeli elfedési görbe avagy az abszolút hallásküszöb alatt maradjon. Erről a jelenségről sok információ található a szakirodalomban, a hangkódolási eljárások általában mégsem használják. Amint látható, kettősség van az időbeli és a frekvenciabeli hallásküszöb-változás között. Mivel a frekvenciafelbontás növelése esetén romlik az időbeli felbontás, és fordítva, a jel pillanatnyi tulajdonságait figyelembe véve

<sup>1</sup> olyan frekvenciasáv, melyen belül lévő frekvencia-összetevőket hallásunk intenzitás alapján összegzi

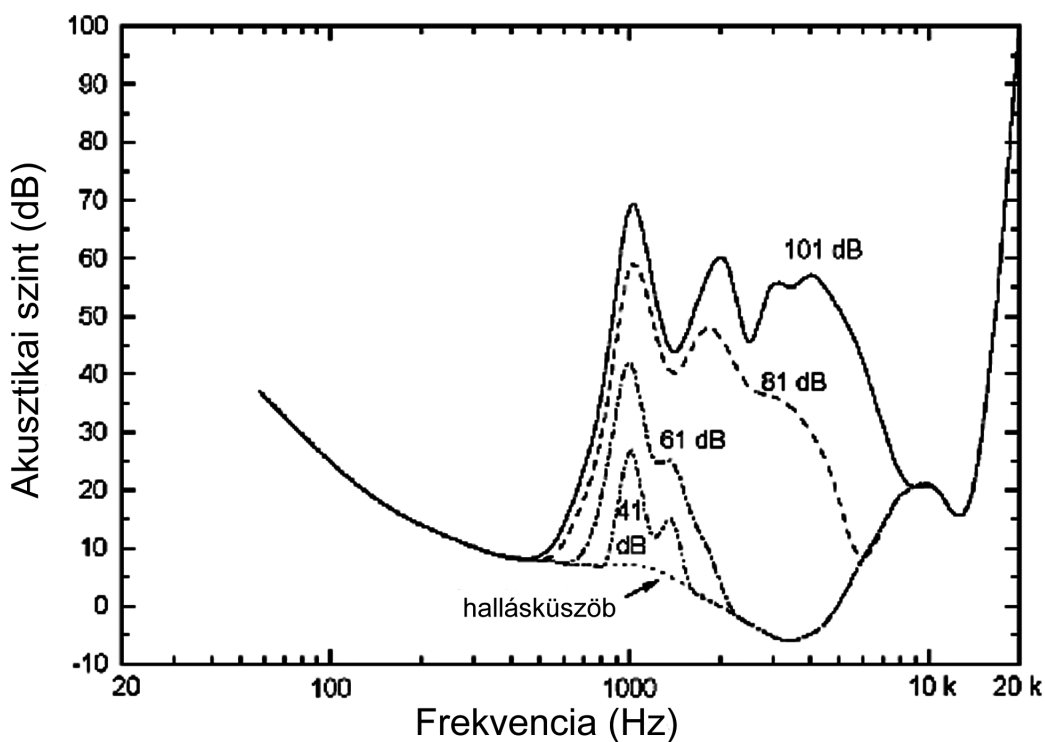


célszerű a frekvencia- ill. az időbeli felbontást előnyben részesíteni. Lassan változó jelekre leghatékonyabban a nagy spektrális felbontású kódolások használhatók, míg tranziens jelekre nagy időfelbontású kódolás alkalmazandó.

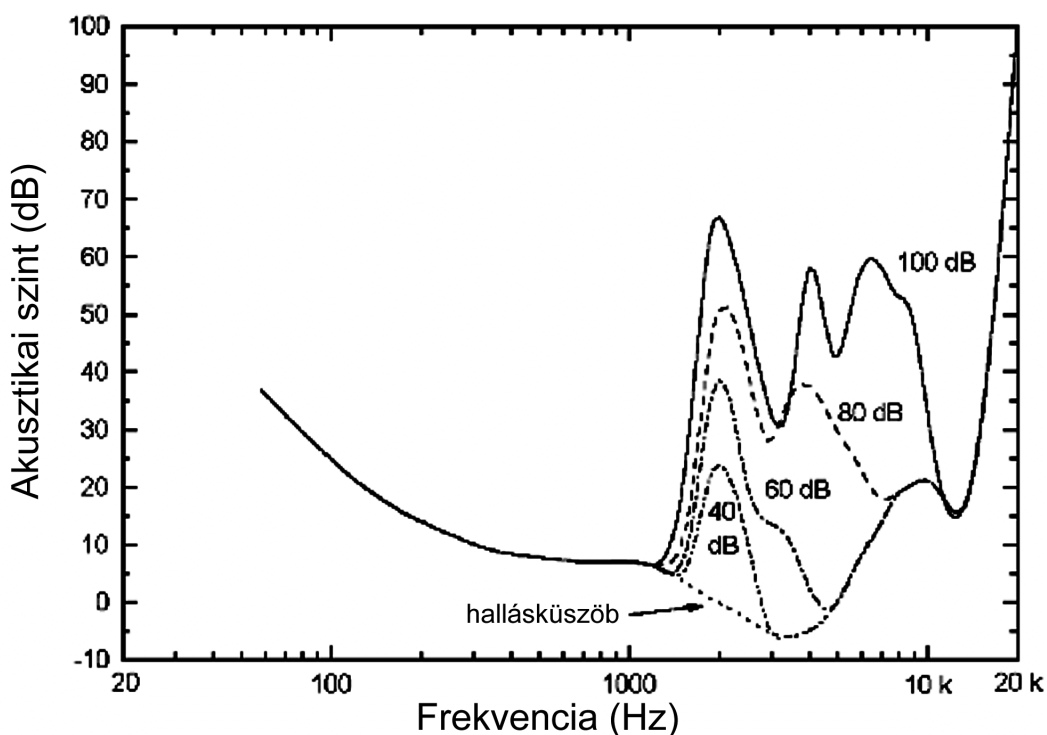
#### **4.1. Spektrális elfedési jelenségek [1]**

Pszichoakusztikus kísérletek kimutatták, hogy az emberi hallás által érzékelt spektrum kritikus sávokra bontható, melyeken belül hallószervünk intenzitásuk szerint összegzi az oda eső jeleket. A sávok frekvenciaátfogása nem egyenlő, viszont minden kritikussáv azonos távolságnak és így azonos számú neuronnak felel meg a baziláris membránon. Ezért amennyiben a bemeneti jelre alkalmazott frekvencia-transzformáció lineáris léptékű spektrumot ad, úgy azt nem-egyenközű sávokra kell bontani, melyek szélessége kritikussáv vagy annál kisebb, majd az egy sávon belüli frekvencia-komponenseket integrálni kell. Az így kapott *gerjesztési függvény* alapján jól vizsgálható hallásunk gerjesztettségi állapota.

Az 4.1. és a 4.2. ábrán spektrális elfedési jelenséget vizsgáló kísérletek eredménye látható. A szubjektív kísérletben a résztvevőknek 1 ill. 2 kHz-es szinuszos maszkoló jel mellett az egyes kritikus sávokra adott zajjal éppen meghallható szintjét kellett meghatározniuk. Az ábrákról leolvasható, hogy a maszkolási hallásküszöb-görbék tartalmaznak egy kis- és nagyfrekvenciás tartományt, amelyek nem különböznek a normál (teljes csönd melletti) hallásküszöbtől, és van egy középtartomány, ahol csökken a hallás érzékenysége. A legnagyobb érzékenységcsökkenés a gerjesztő jel kritikus sávjába esik. A maszkoló jel szintjének növekedtével a középfrekvencia fölötti erősen érzéketlen tartomány egyre szélesedik. Ezzel szemben az alatta lévő tartomány szinte alig változik



4.1. ábra. 1 kHz frekvenciájú szinuszhullám maszkolási görbéi (0 dB SPL = 20  $\mu$ Pa)



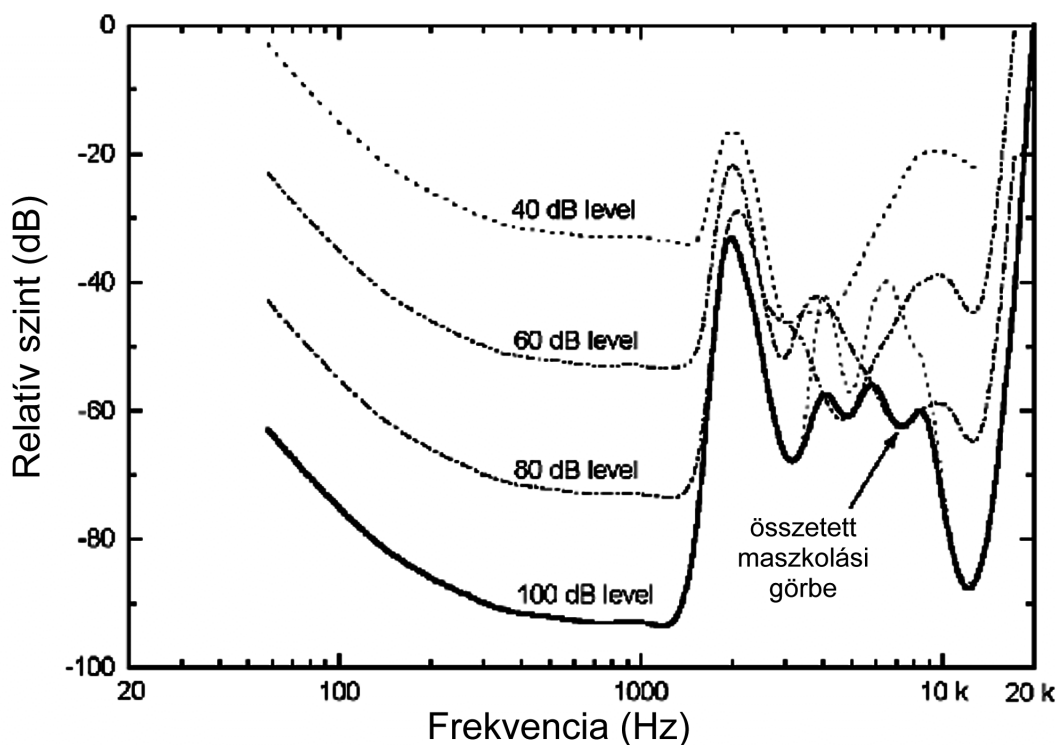
4.2. ábra. 2 kHz frekvenciájú szinuszhullám maszkolási görbéi (0 dB SPL = 20  $\mu$ Pa)

Egy kódolási eljárás során az a cél, hogy a gerjesztő jel függvényében meghatározzuk a pontos maszkolási görbét, majd a jelet úgy kvantáljuk, hogy a jelhez adódó zaj spektrális eloszlása mindvégig a maszkolási görbe alatt

maradjon. A maszkolási görbe meghatározására többféle módszer létezik. Itt most az AC-3 eljárás alkalmazta elvet szeretném bemutatni.

Hogy a maszkolási görbe számítása ne legyen túl bonyolult, el szeretnénk jutni egy olyan függvényhez, mely a maszkoló jel szintjétől és frekvenciájától függetlenül megadja, hogy a maszkoló jel környezetében hogyan módosul a hallásküszöb. Ezt az univerzális függvényt *elterülési függvénynek* nevezzük.

Először a maszkolási görbéket vonatkoztatjuk a maszkoló jel szintjére (4.3. ábra). Majd a görbék minimumából (semmilyen jelszintnél se lépjük túl a tényleges maszkolási görbét) meghatározzuk az *összetett maszkolási görbét* (4.3. ábra vastaggal rajzolt görbéje).



4.3. ábra. 2 kHz-es szinusz jel akusztikai szintjére vonatkoztatott maszkolási görbék  
(0 dB SPL = 20  $\mu$ Pa)

Végül az összes frekvencia összetett maszkolási görbéja alapján képezzük az elterülési függvényt. Ezt úgy számoljuk, hogy egy olyan frekvenciaskálán rajzoljuk meg az egyes összetett maszkolási görbéket, melynek a skálázása relatív kritikussáv-távolságú, és a 0 a maszkoló jel frekvenciáját jelöli. A legkedvezőtlenebb esetet feltételezve képezzük a függvények minimumát, és így



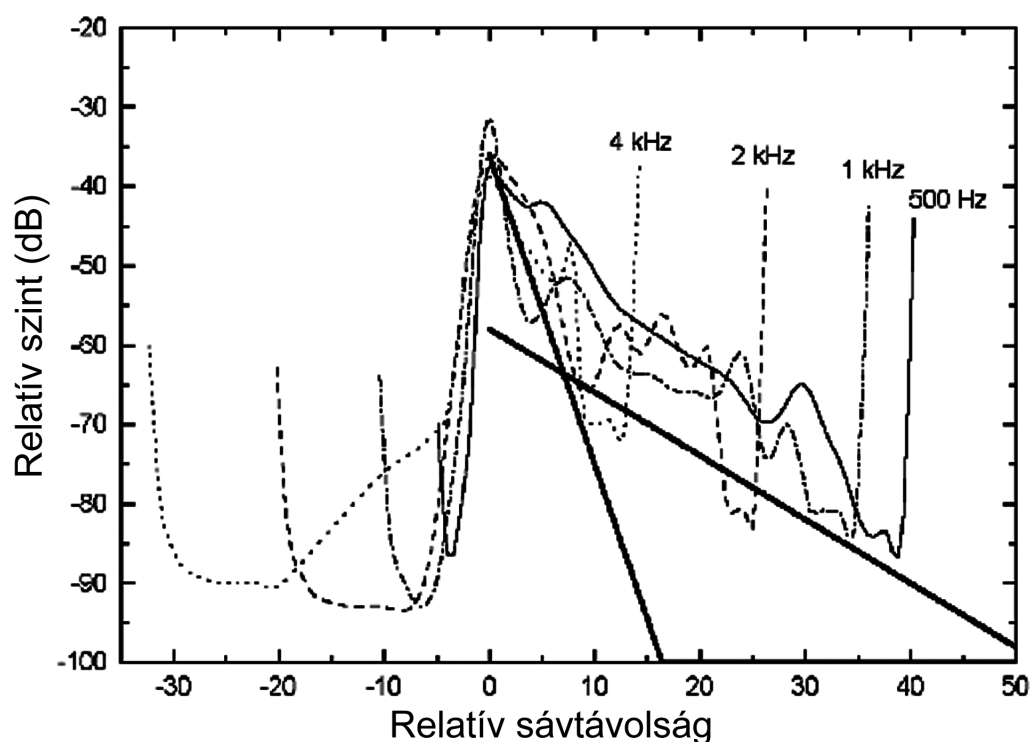
kapjuk az elterülési függvényt. Azt a függvényt, amely valamilyen egyszerű függvényalakkal közelíti az elterülési függvényt, *prototípus maszkolási görbének* nevezzük (4.4 ábra).

**Hallásküszöb:** a legkisebb, még érzékelhető hangnyomásszintek a frekvencia függvényében

**Maszkolási görbe:** egy adott jel hatására módosult hallásküszöb-görbe

**Összetett maszkolási görbe:** egy adott frekvenciájú szinusz jel szintfüggetlen maszkolási görbéje

**Elterülési függvény:** egy szinusz jel frekvencia- és jelszint-független maszkolási görbéje



**4.4. ábra. Prototípus maszkolási görbe meghatározása**

megjegyzés: az ábrán látható sávok szélessége fél kritikus sáv

Az AC-3 kódolás prototípus maszkolási görbéje a 4.4. ábrán látható két egyenesből tevődik össze, a maszkolási görbe könnyebb számíthatósága érdekében a kisebb jelentőségű középfrekvencia előtti darabot elhanyagolja.

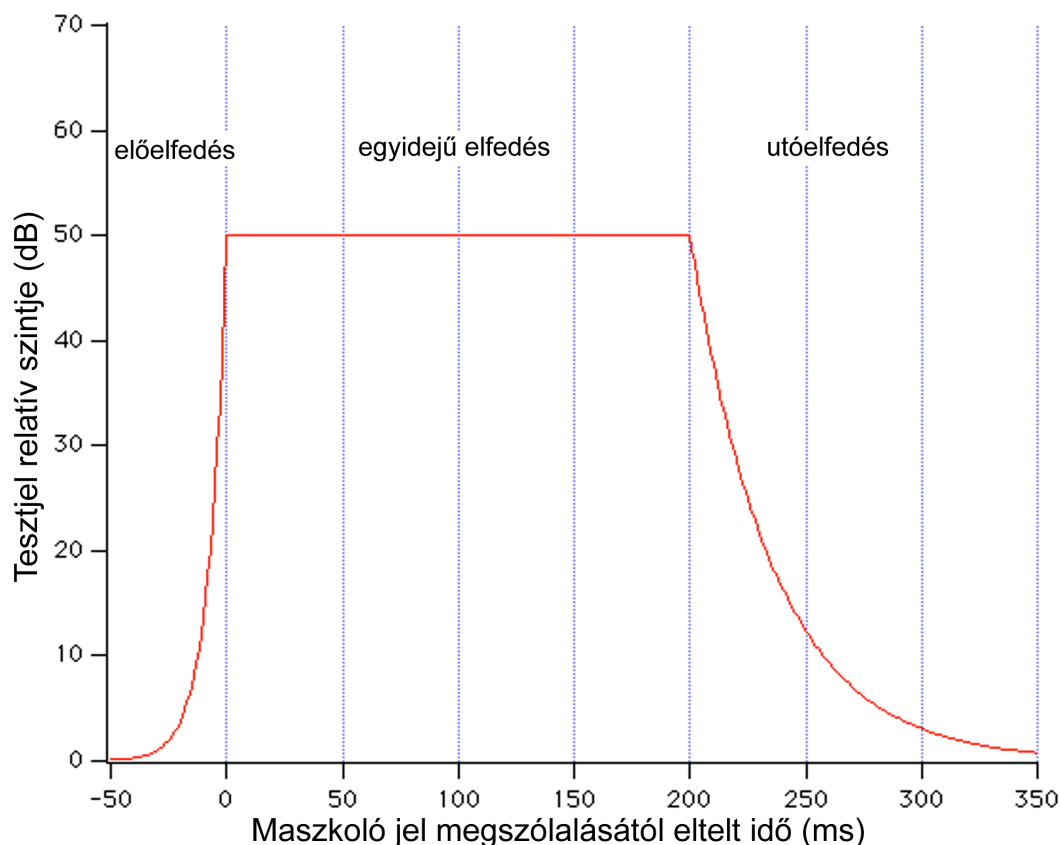
## 4.2. Időbeli elfedési jelenségek [2]

Az időbeli elfedési jelenségeket három csoportba sorolhatjuk:

*Előelfedésnek* nevezzük azt a jelenséget, mikor egy teszttel észlelhetősége az őt követő maszkoló jel hatására csökken. A maszkoló jel befolyásolja, hogy mi az a legkisebb hangerő, amin meghalljuk az őt megelőző tesztjelet, azaz a maszkoló megszólalása „előtt” fejt ki hatását. Ez a jelenség azzal magyarázható, hogy a hallás érzékelése nem pillanatszerű, hanem adott időtartamon belül lefolyó eseményeket összegez.

*Egyidejű elfedésről* beszélünk, mikor a maszkoló és a teszttel együttesen van jelen. Ennek időtartama ilyen módon megegyezik a maszkoló jel kitartási idejével.

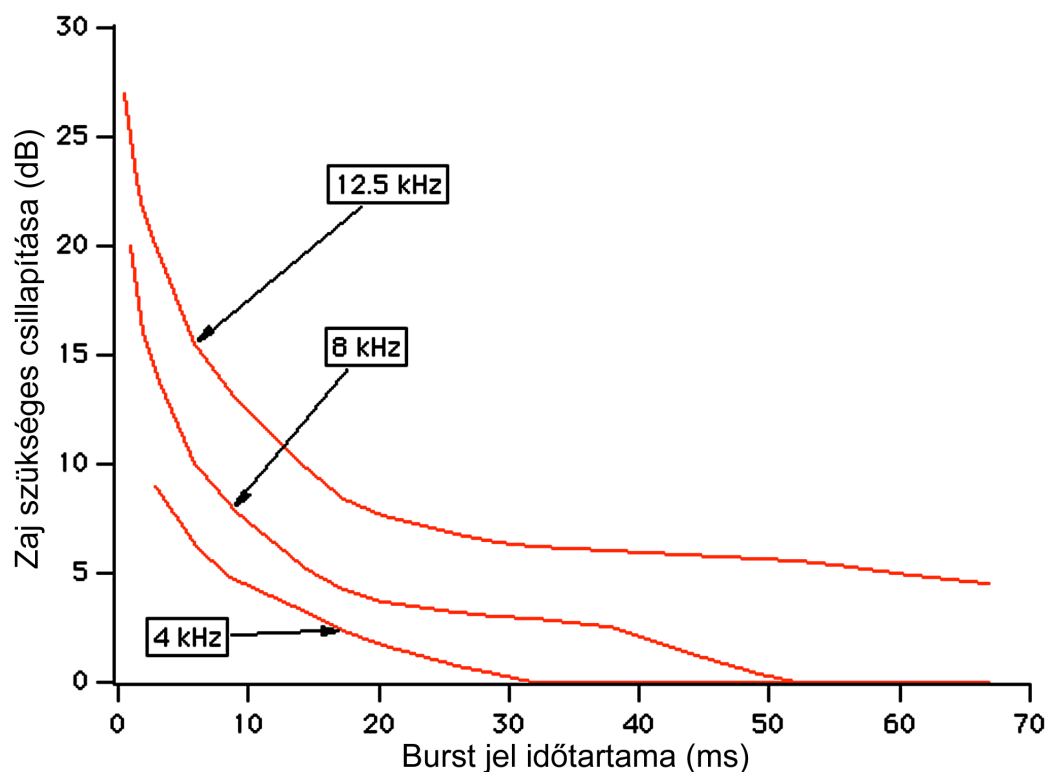
A maszkoló jel elhallgatása utáni időszakban történő eseményt *utóelfedésnek* nevezzük. Ez a jelenség akár egész hosszú lefolyású is lehet. Pl. egy hangos koncert után az ember hallása néha órákig, esetenként egy-két napig is érzéketlenebb, mint normál állapotban.



4.5. ábra.  $t = 0$  időpillanatban megszólaló, 200 ms időtartamú burstjel időbeli elfedési görbéje

Az elfedési jelenségek relatív időtartamát láthatjuk a 4.5. ábrán [3]. Az ábra a tesztjel éppen meghallható szintjét ábrázolja egy relatív hosszan tartó (200 ms) maszkoló jel mellett, arra az esetre vonatkoztatva, mikor nincs maszkoló jel. A maszkoló jel megszólalása  $t=0$  időpillanatban történik. Az 4.5. ábráról leolvasható, hogy az előelfedés nagy hatású időtartama rövid, kb. 10 ms. A maszkoló jel megszólalásától távolodva rohamosan csökken az elfedés mértéke. Ebből azt a következtetést lehet levonni, hogy a legkritikusabb hangjelek esetében 3 ms nagyságrendű időbeli felbontásra van szükség ahhoz, hogy az adott időtartam alatt nem változó kvantálási zaj szintje az előelfedés miatt megemelkedett hallásküszöb alatt maradjon. Az utóelfedés okozta hallásküszöb emelkedés ezzel szemben lényegesen hosszabb, kb. 100 ms, és mintegy 200 ms-nál van teljesen vége. Ez az eredmény nem meglepő, figyelembe véve, hogy az emberi hallás integráló hatását általánosan 200 ms-ra becslik.

A 4.6. ábrán a hallás különböző frekvenciájú és időtartamú jelekre adott reagálása látható. A meghallgatási kísérlet során burstjeleket kódoltak egy fix blokkméretű (11.6 ms hosszúságú) transzformációs kódolóval 128 kbit/s-ra.



4.6. ábra. 128 kbit/s-os, 11.6 ms blokkméretű transzformációs kódoló kvantálási zajának szükséges csillapítása, hogy azt ne lehessen meghallani

A 11.6 ms-os blokkméret elegendő frekvenciafelbontást biztosít kvázistacionárius jelekre, de tranziens esetén akár 8 ms-al korábban megjelenő előzajt is okozhat. A kísérleti elrendezésben egy állítható osztóval lehetett a kvantálási torzítást csillapítani, míg az eredeti burstjel szintje változatlan maradt. A 4.6. ábrán a zajjal éppen meg nem hallhatóságához szükséges csillapításnak a mértéke a burstjel hosszúságának függvényében látható. Megjegyzendő, hogy az előelfedés mértéke erőteljesen függ a maszkoló jel hosszától és frekvenciájától. Az ábrán látható legrosszabb esetben egy 6 periódusidő hosszú 12.5 kHz frekvenciájú jel esetén 27 dB csillapítás szükséges, hogy a zajt ne lehessen meghallani. Fix blokkméretű kóder esetén 27 dB-es kvantálási-zaj csillapítás csak jelentős mértékű bitsebesség-növeléssel érhető el (kb. 4.5 bit/minta). Amennyiben ezt a jelet 3 ms blokkhosszúsággal kódolnánk, egyáltalán nem volna többletbitre szükség.

A fenti gondolatmenetből világosan kiderül, hogy a spektrális elfedési jelenségek kizárólagos használata nem elegendő. Adott bitsebesség mellett a lehető legjobb szubjektív minőség eléréséhez az időbeli elfedési jelenségeket is figyelembe kell venni. Ez hatékonyan úgy valósítható meg, ha a kóder időbeli elfedési modellt használva változtatja a transzformációs blokk méretét. Lassan változó jelekre hosszú, tranziens jelekre rövid blokkméretű szűrőbank használandó.

### 4.3. Bitkiosztás [1, 4]

Bitkiosztásnak nevezzük azt a folyamatot, amikor véges számú  $B$  bitet osztunk szét  $M$  darab minta között úgy, hogy a kiosztás által meghatározott bitszámmra kvantált minták torzulása a lehető legkevésbé legyen hallható. Eredménye egy bithozzárendelés  $b(k)$ ,  $k=0, 1, \dots, M-1$ , mely megadja a  $k$ . minta szóhosszúságát (bitszámát).

$$\sum_{k=0}^{M-1} b(k) = B$$

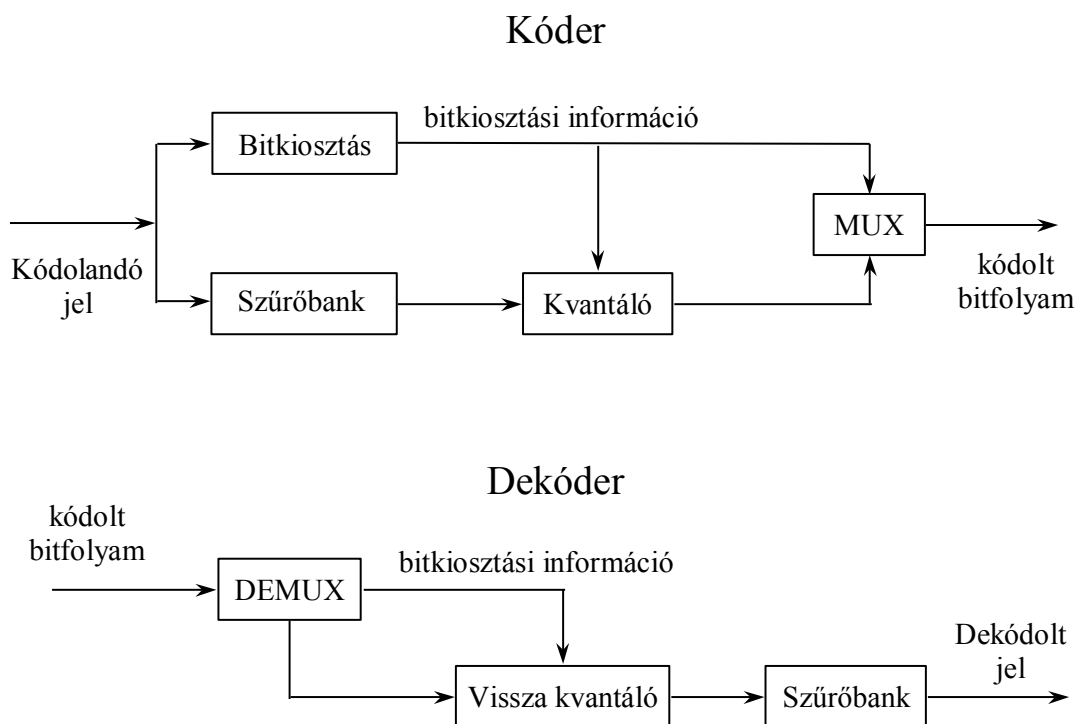
$B$ : a minták között szétosztandó bitek száma:

$B = \text{bitsebesség (kbit/s)} * \text{transzformációs blokk hossza (ms)} - \text{segédinformációk}$

segédinformációk: fejléc, bitfolyam-információk, járulékos adatok, stb.

A dekóderben a kódolt információ pontos visszanyerésének érdekében valamilyen módon elő kell, hogy álljon ugyanaz a bitkiosztás, mint ami szerint a kóder továbbította a mintákat. Erre két alapvető módszer létezik:

Egy *előrefelé adaptív* kóder (4.7. ábra. ) kiszámolja a legmegfelelőbb bitkiosztást, majd azt valamilyen kódolt avagy kódolatlan alakban segédinformációként továbbítja a dekóder felé.

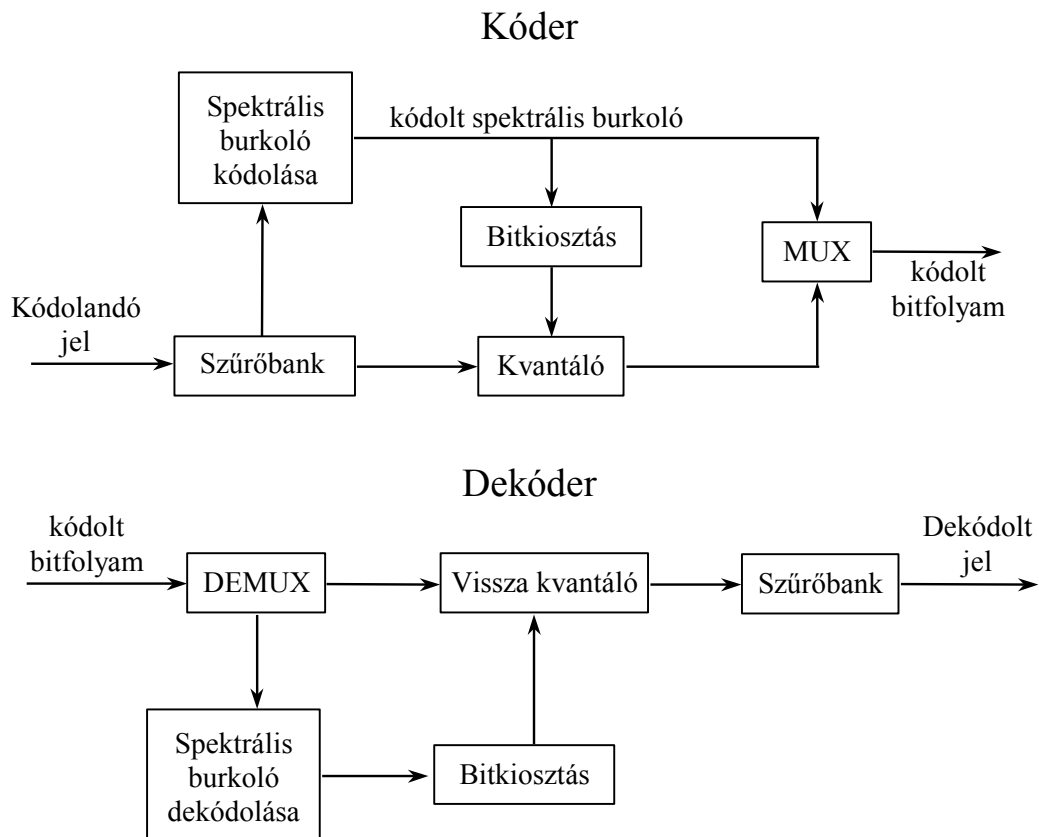


**4.7. ábra. Előrefelé adaptív bitkiosztás**

Elméletileg ez a módszer biztosítja a legpontosabb bitkiosztás lehetőségét, mert a kóder az eredeti jelen számolhat, és szinte bármilyen bonyolultságú lehet. Mivel a műsorszóró rendszerek aszimmetrikusak, nagyságrendekkel több dekóder van a rendszerben, mint kóder, ezért gazdaságilag jó megoldás, ha inkább az a kevés számú kóder kerül sokba, de azzal biztosítják a kiváló minőséget. Egy másik jelentős előnye ennek a módszernek, hogy mivel a pszichoakusztikus modellt csak a kóder tartalmazza, ezért annak módosítása, a technológia fejlődésével történő javulása nincsen semmilyen hatással a meglévő

dekóderek működésére. Ez pedig jelentősen megnövelheti egy kódolási eljárás használatának élettartamát. Mindezen kedvező tulajdonságai mellett hatékonyságának gyakorlati korlátja van, nevezetesen, hogy a rendelkezésre álló bitsebesség egy részét elhasználja a bitkiosztás továbbítása, amely alacsony bitsebességeknél jelentős arányú lehet a teljes bitsebességhez képest. Pl. az ISO MPEG-1 Audio 2-es rétegű (layer II) kóder előre adaptív típusú kóder. A továbbított bitkiosztási információja, mely 24 ms-os idő- és 750 Hz-es frekvenciafelbontású, kb. 4 kbit/s-ot használ el csatornánként a teljes bitsebességből. Tranziens jelek alatt előnyös volna a bitkiosztást valamivel finomabb időfelbontással továbbítani, de ez természetesen jelentősen nagyobb bitkiosztási adatsebességet vonna maga után (2-es rétegű kóder esetén, 8 ms-os időfelbontású bitkiosztás már 12 kbit/s/csatornát igényelne). Kvázi-stacionárius jelek esetén ezzel szemben valamivel nagyobb frekvenciafelbontás volna szükséges. Pl. egy olyan jel esetén, mely minden 750 Hz-es sávban tartalmaz egy spektrumvonalat, minden sáv számára foglalna biteket. Amennyiben a frekvenciafelbontás nagyobb lenne, úgy az egyes spektrumvonalak közötti frekvenciákra nem kéne biteket foglalni és ezzel hatékonyabb lenne a bitsebesség-csökkentés. Sajnos azonban mind a frekvencia-, mind az időfelbontás növelése jelentős továbbítandó többletinformációval járna. Ezért elméletileg nagyon előnyös ez a módszer, gyakorlatilag azonban komoly korlátokkal bír alacsony bitsebességek esetén.

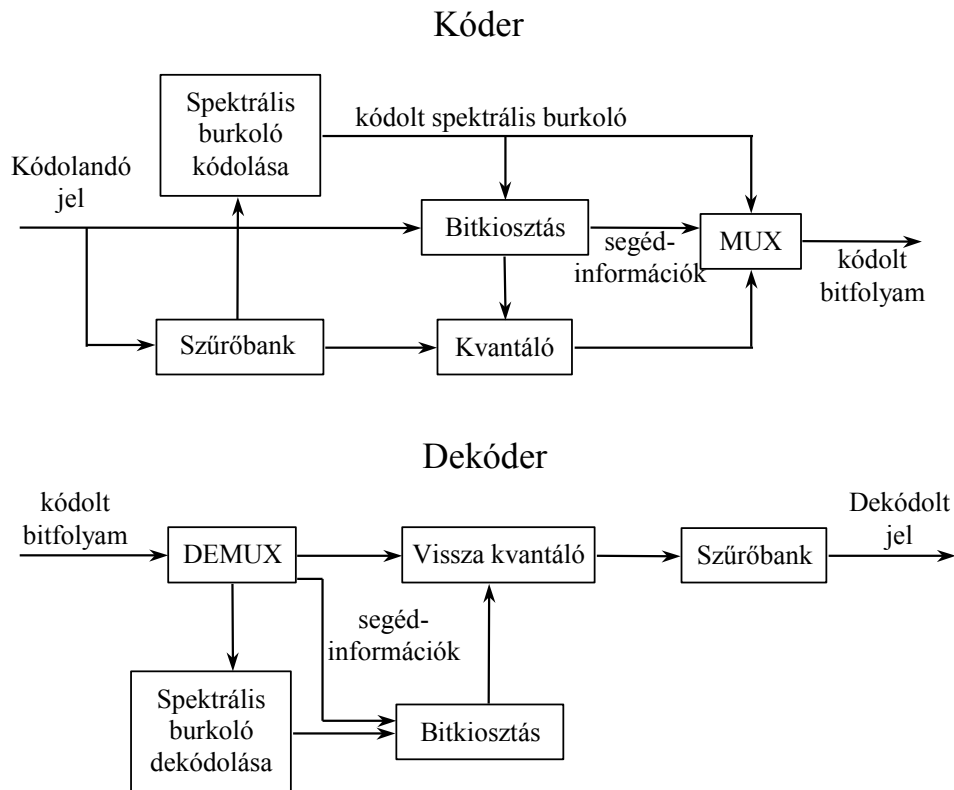
A *visszafelé adaptív* kódolási módszer lényege, hogy a bitkiosztás a kódolt audio adatból számolandó (4.8. ábra).



4.8. ábra. Visszafelé adaptív bitkiosztás

Ennek a módszernek az az előnye, hogy semmit nem használ el a bitsebességből a bitkiosztási információ, így az összes bit a minták kódolására fordítható. A bitkiosztás lehet mind frekvenciában, mind időben nagy felbontású a jel szerkezetétől függően. Emiatt a visszafelé adaptív rendszerek nagyon hatékonyak átviteli sebesség szempontjából nézve. Hátrányuk abból a tényből következik, hogy a dekóderben a bitkiosztást a bitfolyamban található információból kell számolni. Így a bitkiosztást már csak korlátozott pontosságú információkból nyerjük, és ezáltal kisebb-nagyobb hibákat vétünk. Másrészt az alkalmazott bitkiosztási algoritmus nem lehet túl bonyolult, mert az nagyon költségessé tenné a dekódereket. Ráadásul a dekóderek piacra kerülése után többet már nem lehet változtatni, finomítani a pszichoakusztikus modellt.

Amint az a fentiekből kitűnik, önmagában egyik módszer sem üdvöztető. A kettő ötvözetéből készült hibrid módszer a visszafelé adaptív bitkiosztási algoritmus hátrányainak nagyobb részét kiküszöböli (4.9. ábra. ).



**4.9. ábra. Hibrid előrefelé/visszafelé adaptív bitkiosztás**

Ez a módszer egy visszafelé adaptív modellből indul ki, melyhez opcionálisan előrecsatolható kétfajta járulékos információ, akár a hallásmodell finomítására, akár közvetlen bithozzárendelés megadására. Ezzel lehetőség nyílik a dekóderben egy nem túl bonyolult, de pszichoakusztikus modellen alapuló és nagyjátlagban elég pontos bitkiosztás számolására egyedül csak a hallásmodell-paraméterek időközönkénti módosításával, extrém körülmények között pedig a számolt eredmény közvetlenül módosítható a kóder oldaláról.

A bitkiosztás magvát képező rutin pszichoakusztikus modellje egy sor feltételezést tesz a jelek maszkoló tulajdonságát illetően. A modell egyes paraméterei változtathatók. Így a dekóderben használt modell finomabb részleteit – a bitfolyamban továbbítva – a kóder közvetlenül beállíthatja. A kóder akármilyen pontos modellt használhat, és annak eredményeit összevetheti a dekóderben használt alap-moddellel. Ha modellparaméterek állítgatásával a valóságot jobban közelítő eredmény érhető el, akkor ezzel élhet a kóder. Amennyiben ez nem vezetne elég jó eredményre, úgy módjában áll közvetlenül bithozzárendelés-értékek módosítására is.



#### 4.4. MPEG Audio

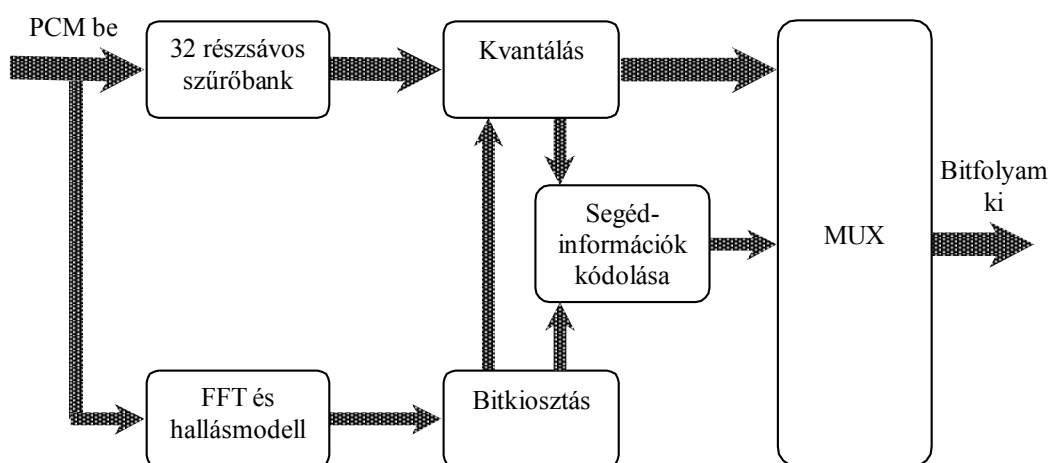
Az MPEG szó az 1988-ban alakult ISO/IEC Motion Picture Experts Group rövidítése. Ennek a munkacsoportnak az volt a feladata, hogy egy egységes digitális kép és hang redundancia-csökkentő rendszer-családot dolgozzon ki. A hangkódolási eljárásokkal az audio csoport foglalkozott. Mivel ekkor már nagyon sokfajta hangkódolási eljárás létezett, ezért nem egy új eljárást hoztak létre, hanem a legkülönbözőbb szempontokat figyelembe véve a már meglévők közül válogattak. Munkájuk eredménye az MPEG-1 Audio szabvány [5], majd annak továbbfejlesztett változata, az MPEG-2 Audio szabvány [6].

##### MPEG-1 Audio [5]

Az MPEG-1 Audio-ban három kódolási eljárást definiáltak: 1-es, 2-es és 3-as réteget. (Layer-1, -2, -3), mivel célszerűnek láttak eltérő bonyolultságú és minőséget szolgáltató kódolásokat bevezetni. Az 1-es-től a 3-as felé haladva a kóder (s a dekóder) bonyolultsága és hatékonysága egyre nő. Az 1-es réteg a 2-es leegyszerűsített változata, így annak ismertetését itt mellőzöm.

##### 2-es réteg

Ez a kódolás tisztán a MUSICAM (Masking pattern adapted Universal Sub-band Integrated Coding And Multiplexing) eljárás. A kóder leegyszerűsített blokkvázlata a 4.10. ábra. ábrán látható.



4.10. ábra. MPEG Audio 2-es rétegű kóder blokkvázlata

A bemeneti jel 32, 44.1 vagy 48 kHz mintavételi frekvenciájú PCM-mintasorozat lehet. A jel kétfelé ágazik. Az egyik ágon haladva, mely a

maszkolási görbét számolja, egy 1024 pontos frekvencia-transzformáció után 32 részsávra bontjuk a jelet, majd kiszámoljuk az egyes sávokba eső energiát. Egy pszichoakusztikus modell segítségével minden sávra kiszámoljuk, hogy mennyi a megengedhető kvantálási zaj. A másik ágon egy szűrőbank 32 azonos szélességű részsávra bontja a jelet. Az egyes részsávok egymást követő (idő-) mintái párhuzamosan haladva azonos feldolgozási lépéseken mennek keresztül. Minden részsávból 36 mintát véve kapunk egy keretet. Ennek lejátszási időtartama 48 kHz-en 24 ms, ez az újraszinkronizálási egység. A keretet 3 blokkra osztjuk. Mindhárom blokk egy-egy részsávjára számolunk egy léptéktényezőt (előre meghatározott értékek közül legkisebb olyan szám, mely mind a 12 mintánál nagyobb), majd ezzel leosztva a mintákat normalizált alakhoz jutunk. Így részsávonként 3 skálázási tényezőnk és 36 mintánk van egy keretben. Ezután a 36 mintát a hallásmodell által kiszámolt szükséges jel/zaj viszony és a kimeneti bitsebesség függvényében adott bitszámra kvantáljuk. Mind a 36 minta azonos bitszámra kódolódik. A kódolt bitfolyamban ezután a keret felépítése a következő (lásd még 4.11. ábra.):

Fejléc: szinkronszó, azonosító, rétegszám, bitsebesség, mintavételi frekvencia, stb.

CRC hibadetektáló kód

Bitkiosztást jelző információ (adott részsáv mintáit hány biten továbbítjuk) részsávonként és csatornánként

Léptéktényező választási információ: amennyiben egy részsáv 3 léptéktényezője közül 2 vagy 3 egymást követő azonos, úgy elég csak egy vagy kettő léptéktényezőt továbbítani. Ennek jelzésére szolgál a léptéktényező választási információ (részsávonként és csatornánként 2 bit)

Léptéktényező: részsávonként és csatornánként 3 vagy kevesebb 6 bites léptéktényező a normalizált részsáv minták visszaalakításához

Részsáv minták kvantálva: egymás után 3 blokk, egyenként 4\*3 mintával, 3 mintás egységenként a 32 részsáv mintái csatornánként egymásba szőve

## Járulékos adatok

### 3-as réteg

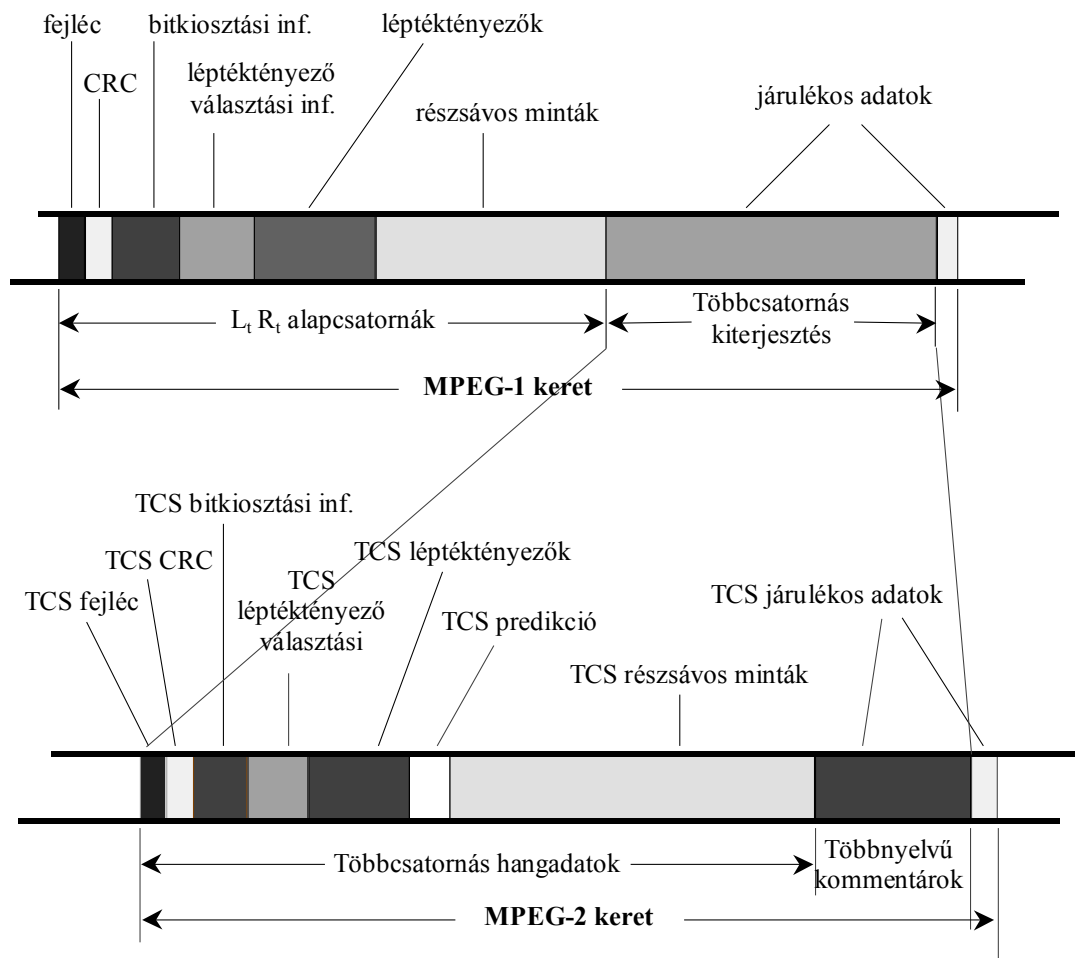
Ez a réteg jóval bonyolultabb, mint a másik kettő, az ASPEC (Adaptive Spectral Perceptual Entropy Coding) és a MUSICAM eljárások ötvözéséből készült. Ennek eredményeképpen az általa elérhető kódolási hatékonyság is jóval nagyobb. Egy 3-as rétegű kóder a kvázi CD-minőséget kb.128 kbit/s/csatorna bitsebesség mellett képes teljesíteni, míg a 2-es rétegnek ehhez kb.192 kbit/s/csatornára van szüksége [7]. A 2-es réteg legnagyobb hibája, hogy a bitkiosztás időbeli és frekvenciafelbontása nem elegendő. A 3-as réteg ezt kijavítja. A kóder nagyon hasonlít a 2-es rétegére, csak itt a részsávos mintákat egy 50%-os átlapolású módosított diszkrét koszinusz transzformációnak (MDCT) vetik alá. Tranziens jelek esetén a transzformációs ablak 12 minta nagyságú, stacioner jeleknél 36 minta. A 36 mintából 18 spektrumvonal lesz, így a bitkiosztást nem 32 (2-es réteg), hanem 576 frekvencián lehet számolni 24 ms-os felbontással, tranziens esetén pedig 3 darab 8 ms idejű, 192 frekvenciát tartalmazó blokkra. Ez már elegendően nagy felbontás ahhoz, hogy hatékonyan lehessen a biteket kiosztani. Ezen felül a 3-as réteg nem-egyenközű kvantálóval rendelkezik, a kvantált hangmintákat entrópia-kódolással továbbítják, és tartalmaz egy úgynevezett *bittartályt*, mely lehetővé teszi, hogy a különösen magas bitigényű keretek adatát az előtte és mögötte elhelyezkedő keretekre szét lehessen osztani, amennyiben azok kereteiben marad szabad bit.

### Kódolási módok

A szabvány 4 lehetséges kódolási módot ír elő: egycsatornás, kétcsatornás, sztereó és kapcsolt sztereó. A kétcsatornás és a sztereó közötti különbség az, hogy a sztereó módban a két csatorna közös bitkészleten osztozik, a csatornák pillanatnyi bitigényétől függően hol az egyik, hol a másik kap több bitet. Kétcsatornás üzemmódban a bitkészletet felezik, és mindkét csatorna ugyanannyi bitet kap. Ebben az üzemmódban a két csatornát egymástól függetlenül lehet kódolni.

A kapcsolt sztereó üzemmód 1-es és 2-es réteg esetén intenzitásos sztereó kódolást jelent, 3-as réteg esetén intenzitásos és MS sztereó mód között

választhatunk. Az emberi hallásnak van egy olyan tulajdonsága, hogy a kisfrekvenciás jeleket pillanatnyi fázisuk alapján lokalizálja, míg a frekvencia növekedtével egyre inkább csak a két fülbe jutó intenzitás nagysága számít. Ezt használja ki az intenzitációs sztereó kódolás: egy megadott frekvenciától kezdve, a két csatorna normalizált részsávós mintáit átlagolva egy közös csatornát képez, majd ezt továbbítja. Dekódolás során mindkét csatorna a közös csatorna mintáit kapja meg. Így a közös csatorna alsó frekvenciája fölött feleannyi információt kell továbbítani. Mivel az egyes csatornák léptéktényezői továbbra is megmaradnak, ezért a csatornák energiája megmarad, csak a fázisinformáció vesz el. A 3-as rétegben megengedett MS sztereó mód azt jelenti, hogy nem a bal és a jobb csatorna mintáit kódolják, hanem  $bal_{MS}=bal+jobb$  és a  $jobb_{MS}=bal-jobb$  csatornákat képeznek, majd ezek szerint kódolnak. A dekóder előállítja a bitfolyamból a  $bal_{MS}$  és  $jobb_{MS}$  csatornákat, majd azokból a balt és a jobbot. Ezt az üzemmódot arra az esetre találták ki, amikor egy Dolby Surround jelet kívánunk MPEG kódolással továbbítani. Attól függően, hogy melyik csatornában található a domináns jel (bal, jobb,  $bal_{MS}$ ,  $jobb_{MS}$ ), a kóder bal és jobb, avagy  $bal_{MS}$  és  $jobb_{MS}$  szerinti kódolásra dönt. Ez keretenként adható meg. Ráadásul 3-as réteg esetén az MS sztereó és az intenzitációs sztereó egyidejűleg is használható, ilyenkor a frekvenciatartomány alsó részében MS sztereó módban kódolnak, a közös csatorna alsó frekvenciájától kezdve pedig intenzitációs sztereó módban.



4.11. ábra. Az MPEG Audio keretszerkezete

## MPEG-2 Audio [6]

Ez a szabvány ugyanazt a kóder - dekóder családot használja, mint az MPEG-1: 1-es, 2-es és 3-as réteget. Az MPEG-2 újdonságai:

**alacsony mintavételi frekvenciás kiterjesztés:** nagyon alacsony bitsebességek esetén hatékonyabb a kódolás, ha a bemeneti mintavételi frekvenciát csökkentik, nem pedig az átviendő részsávok számát. Az új frekvenciák a már meglévőknek a felezett értékei, tehát 16, 22.05 és 24 kHz.

**többcsatornás kiterjesztés:** a 3/2-es sztereó hangformátumnak megfelelő csatornákat és egy kisfrekvenciás effektcsatornát lehet kódolni

**többnyelvű kiterjesztés:** lehetőség van 7 további csatorna továbbítására a többnyelvű kommentárok érdekében

## **Kompatibilitás**

Az MPEG-2 Audio-val szemben támasztott legfőbb követelmény az, hogy az MPEG-1 Audio-val kompatibilis legyen. Ennek három értelemben kell teljesülnie:

**lefelé kompatibilis:** a 3/2 módban kódolt hanganyag kisebb hangsugárzószám esetén is használható legyen. Mátrixegyenletek adják meg, hogy a kódolt csatornák mely keveréséből állnak elő a kimeneti csatornák.

**visszafelé kompatibilis:** egy MPEG-2-ben kódolt bitfolyam teljes audió információját ki tudja nyerni egy MPEG-1 dekóder, természetesen csak a kéthangsugárzás kiépítettségre vonatkozóan. Ez a bitfolyam speciális kialakításával és mátrixolósos technikával érhető el.

**előrefelé kompatibilis:** egy MPEG-1-ben kódolt bitfolyamot az MPEG-2 dekódernek le kell tudnia játszani

## **Bitfolyam-kialakítás**

Az MPEG-2 bitfolyam kialakítása a 4.11. ábra. án látható.

Az alapötlet az, hogy minden új információt, amely egy MPEG-2 dekódernek szükséges, az MPEG-1 bitfolyam „járulékos adatok” mezőjébe kell tenni. A két alapsatornába – bal<sub>t</sub> és jobb<sub>t</sub> – bele kell keverni az összes csatornát, hogy az MPEG-1 dekóder minden információt megkapjon, három tetszőlegesen kiválasztott csatornát pedig a járulékos adatok mezőbe helyezünk. Hogy melyik hármat, azt a jelek pillanatnyi állapota szerint választjuk meg.

## **Kiegészítő bitfolyam**

Előfordulhat, hogy az MPEG-1 által előírt maximális bitsebesség (1-es réteg: 448 kbit/s, 2-es réteg: 384 kbit/s, 3-as réteg: 320 kbit/s) nem elegendő a többcsatornás információ átvitelére. Erre az esetre definiáltak egy úgynevezett *kiegészítő bitfolyamot*. Minden olyan adatot, amely nem fér be az MPEG-1 keretbe, a kiegészítő bitfolyamban továbbítunk. Erre van lehetőség, mivel az MPEG kódolások definiálják, hogy milyen módon lehet különböző bitfolyamokat egybemultiplexálni. Az MPEG-2 bitfolyamban néhány járulékos paraméterrel jelzik a kiegészítő bitfolyam jelenlétét.

### **Kompatibilis mátrixolás**

A bemeneti csatornák leképzését a továbbítandó csatornákra mátrixegyenletek adják meg. Egy lehetséges leképzés:

$$B_t = B + c_0 K + c_1 BH$$

$$J_t = J + c_0 K + c_1 JH$$

$$T_0 = K$$

$$T_1 = BH$$

$$T_2 = JH$$

1. egyenlet. MPEG-2 mátrixegyenlet  $T_0$ ,  $T_1$ ,  $T_2$  egy lehetséges megadásával

A mátrixolás során két probléma merül fel:

Ha a domináns jel pl. a középcsatornában van, akkor  $B_t$  csatorna megengedhető kvantálási zaját a középcsatorna jeléhez számoljuk. Dekódolás és az inverz mátrix végrehajtása után a maszkoló komponens eltűnik a bal csatornából, de a kvantálási zaj ott marad és hallhatóvá válik. Ezért nagyon fontos a maszkolási hiba-mentes mátrixolás kidolgozása. Ennek egy lehetséges megoldása, hogy a  $B_t$  és a  $J_t$  bitkiosztását kizárólag a bal és a jobb csatorna maszkolási görbéinek alapján számítják. A mátrixba pedig a 3 külön csatorna kódolt - dekódolt változatát vezetjük, így dekódolás és visszamátrixolás után a bal csatornából a többi csatorna kvantálási zaja teljesen eltűnik, a megmaradó kvantálási zaj pedig az eredeti bal csatornához van igazítva, tehát megfelelő bitsebesség mellett a jel azt elfedi.

Nem lehet optimális a tömörítés, mivel a 3 külön is továbbított csatornát kétszer visszük át: egyszer  $B_t$  ill.  $J_t$  csatornában, egyszer pedig külön.

A mátrixolás, ami a visszafelé kompatibilitás miatt szükséges, olyan szigorú követelmény, hogy annak betartása jelentősen csökkenti a kódolás hatékonyságát. Ezért később bevezették az MPEG-2 audio visszafelé nem kompatibilis változatát.

### **Több csatorna jelenlétének kihasználása**

#### **Csatornaváltás $T_0$ , $T_1$ , $T_2$ megválasztásánál**

Keretenként szabadon adható meg, hogy  $T_0$ ,  $T_1$ ,  $T_2$  továbbítási csatornában mit viszünk át. Arra is lehetőség nyílik, hogy részsávcsoporthoz

külön-külön adjunk meg definíciót. Amennyiben ezt tesszük, növeljük a szükséges segédinformációkat, de esetleg hatékonyabban lehet tömöríteni a csatornákat.

### **Dinamikus áthallás**

Hasonlóan az MPEG-1 által használt intenzitásos sztereó üzemmódhoz, itt is lehetőség nyílik közös nagyfrekvenciás csatorna(-ák) részsávcsopontonkénti kialakítására. Ez úgy történik, hogy egy adott csatorna valamely részsávcsoportjának normalizált mintáinak továbbítása helyett azt jelzik, hogy melyik már dekódolt csatorna normalizált mintáit kell használni. Az így kapott mintát a saját léptéktényezővel beszorozva kapjuk meg az adott csatorna hangmintáját. Amennyiben egy adott csatorna adott részsávcsoportja dinamikus áthallást használ, úgy annak sem bitkiosztási információja, sem kvantált mintái nem kerülnek továbbításra.

### **Középcsatorna fantomkódolása**

Szubjektív tesztekkel kimutatták, hogy nem romlik a középcsatorna lokalizálhatósága, ha csak a 9 kHz alatti frekvenciatartományt kódoljuk önálló csatornaként, a magasfrekvenciás összetevőket pedig bal és jobb csatornába keverve továbbítjuk. A lehallgatás során a középcsatorna magasfrekvenciás tartománya fantomközépként fog megszólalni. Ezzel jelentős mennyiségű bitet lehet megspórolni, mivel a középcsatorna bitkiosztási információit, léptéktényezőit, kvantált mintáit csak az alsó 12 részsávra kell továbbítani.

### **Adaptív többcsatornás predikció**

Ez az eljárás a csatornák közötti redundanciát igyekszik kihasználni. Amennyiben engedélyezve van, úgy  $T_0$ ,  $T_1$ ,  $T_2$  csatornák mintáit predikálják  $B_t$  és  $J_t$ -ből, és csak a predikciós hibát továbbítják. Segédinformációként továbbítani kell a predikciós együtthatókat is, de ez az eljárás a legtöbb esetben még így is adatmennyiség-csökkenéshez vezet.

### **Többnyelvű kiterjesztés**

Az MPEG-2 bitfolyam járulékos információ mezőjébe hét további csatornát lehet elhelyezni többnyelvű szövegek továbbítására. Ezeket a csatornákat lehet



mind az MPEG-1 bitfolyam alapsatornáinak mintavételi frekvenciájával, mind pedig annak felével továbbítani. Nem lehet azonban kapcsolt sztereó üzemmódot használni. A többnyelvű kiterjesztés csatornáit ugyanolyan réteggel kell kódolni, mint az MPEG-2 bitfolyam  $B_t$  és  $J_t$  csatornáit, de legalább 2-es rétegen.

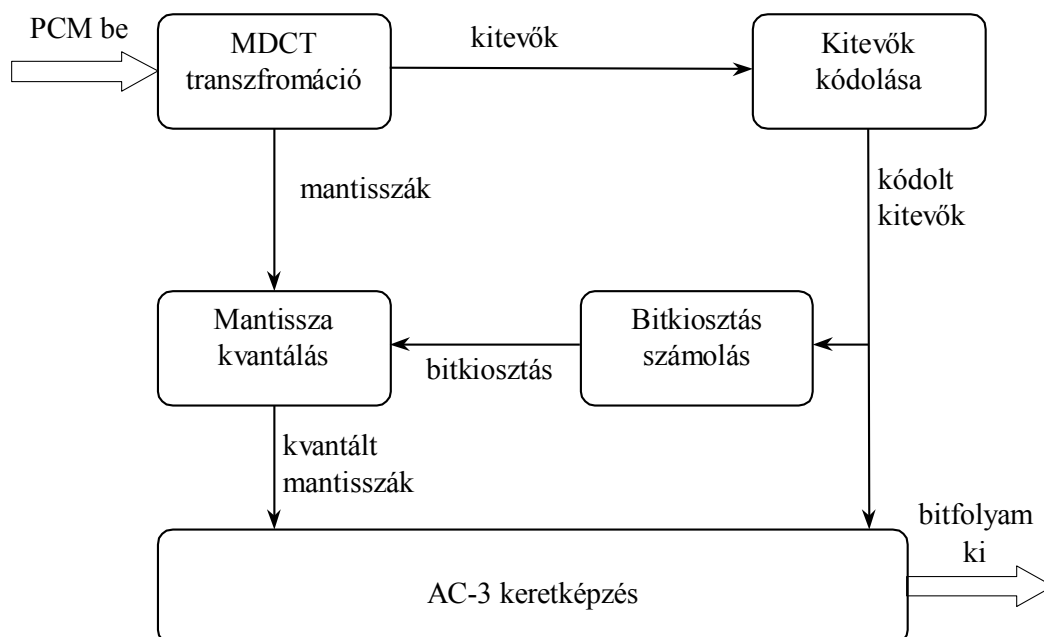
#### 4.5. AC-3

Az AC-3 a Dolby Laboratories által kifejlesztett harmadik pszichoakusztikus kódolás, melyet 1992-ben vezettek be Dolby Digital néven a filmszínházakban. Nagy sikerére való tekintettel előterjesztésre került a nagyfelbontású televízió (HDTV) kísérőhangjaként is. Az ATSC (Advanced Television Systems Committee) 1994-ben szabványosította mind nagyfelbontású, mint normál felbontású TV hangjaként. Az új DVD lemez kísérőhangjaként is a Dolby Digital-t választották NTSC-területek számára (Észak-Amerika és Japán).

Az AC-3, hasonlóan kétcsatornás sztereó elődjéhez, az AC-2-höz, szintén pszichoakusztikus elven működő adaptív transzformációs kódolás, de új, 5.1 csatorna hatékony tömörítésére alkalmas bitkiosztási algoritmust tartalmaz. Hatékonyságát jelentősen növeli, hogy nem kompatibilis elődjeivel, így a bitsebesség-csökkentést maximálisan hozzá tudták a többcsatornás felvételek tulajdonságaihoz igazítani.

#### A kódolás menete

A kóder blokkvázlata a 4.12. ábra. ábrán látható.



4.12. ábra.AC-3 kóder blokkvázlata

Bementi jelként 32, 44.1, 48 kHz-en mintavételezett, maximálisan 5.1 csatornát (3/2-sztereó + kisfrekvenciás effektcsatorna) képes fogadni. A kódolt hanganyagot 1536 PCM-hangminta információját hordozó AC-3 keretekben továbbítják. Ez az újraszinkronizálási egység, 48 kHz-es mintavételi frekvencia esetén 32 ms időtartamú. Az AC-3 keretet 6 audió blokkra osztják föl, egyenként 256 mintányi információtartammal. Az egymást követő transzformációs ablakok 512-PCM mintát 50 %-os átlapolással, MDCT transzformációval 256 spektrumvonallá transzformálnak, és képeznek egy audió blokkot. Amennyiben a bemenő jel tranzienseket tartalmaz, lehetőség van az egy 512 pontos, hosszú transzformáció helyett két egymást követő, 256 mintára alkalmazott rövid transzformációt használni. Ez azt jelenti, hogy stacioner jelek esetén hosszú transzformációt alkalmazva (512 PCM→256 spektrumvonal) 93.8 Hz-es frekvencia- és 5.3 ms-os időfelbontással kódolhatunk, míg tranziens esetén rövid transzformáció mellett (256 PCM→128 spektrumvonal) 187.5 Hz-es a frekvencia- és 2.7 ms-os az időfelbontás. Ezek kiváló értékek.

### Transzformálás

Első lépésként tehát a jelet tranziensvizsgálat után a frekvencia-tartományba transzformáljuk. A használt számábrázolás fractional számábrázolás, ami azt jelenti, hogy minden minta kisebb mint 1. 16 bites

pontosság esetén a 32767 maximális érték  $1 \cdot 2^{-15}$  ( $32767/32768$ , 0.999969482)-t jelent. A frekvencia-együtthatókat lebegőpontos alakra (kitevőre és mantisszára) bontjuk: a kitevő megadja, hogy hány 0-ás van tizedespont és a balról első 1-es előtt, a mantisszát pedig úgy kapjuk, hogy az eredeti frekvencia-együtthatót kitevő bittel balra shifteljük.

Példa 16 bites pontosság esetére:

integer alak (10-es számrendszerben): 5000

fractional alak (2-es számrendszerben): 0.001 0011 1000 1000

kitevő (5 biten): 00010 (=2)

mantissza (16 biten): 0.100 1110 0010 0000 (=20000)

Kitevők kódolása

A kitevő továbbítása 4-fajta kitevő-stratégiával lehetséges:

**újrahasználás:** ilyenkor egyetlen kitevőt sem továbbítunk, a dekóder az előző audió blokk kitevőkészletét fogja használni

**D15:** minden egyes kitevőt továbbítunk

**D25:** csak minden második kitevőt továbbítunk. Dekódolás során két egymást követő kitevő kap meg egy továbbított kitevőt, ezért a kódernek a kitevőket páronként megegyezővé kell tennie.

**D45:** a D25-ös stratégiához hasonlóan, csak minden negyedik kitevőt továbbítunk

A továbbítandó kitevők közül az elsőt (DC komponens) 4 biten továbbítjuk, az összes többi pedig differenciálisan, 2 biten. 3 differenciálisan kódolt továbbítandó kitevőt 5 biten ábrázolunk:

$$\text{kódolt csoportkitevő} = 25 \cdot M[n] + 5 \cdot M[n+1] + M[n+2]$$

$$\text{diff. kitevő}[n] = \text{továbbítandó kitevő}[n] - \text{továbbítandó kitevő}[n-1]$$

$$M[n] = \text{diff. kitevő}[n] + 2$$

### **Maszkolási görbe számolás, kitevők kódolása**

A kitevők adják a spektrum burkolóját, és ez alapján számoljuk a maszkolási görbét, majd később a bitkiosztást. A maszkolási görbét 50 frekvenciasávra (átlagosan fél kritikussáv szélességű) számoljuk. A kitevőket ezután kódolt alakban továbbítjuk a dekóder felé. Mivel az AC-3 kódolás alapján hibrid előre/hátra adaptív bitkiosztást alkalmaz, a dekóder a dekódolt kitevők alapján számolja a bitkiosztást, ezért a kódernek kódolnia, majd dekódolnia kell a kitevőket, és csak ezután lehet számolni belőle a maszkolási görbét.

### **Bitkiosztás számolás**

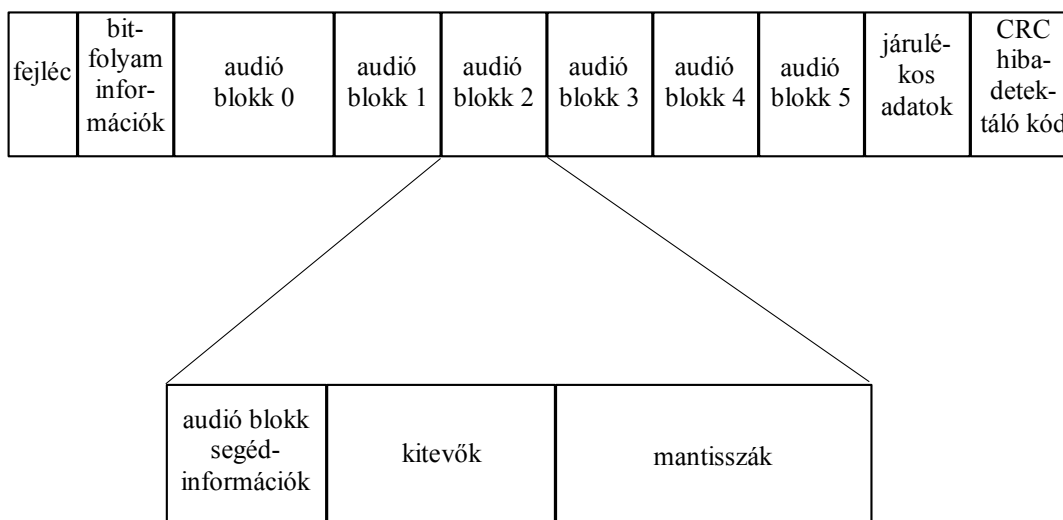
Az 50 pontos maszkolási görbét nulladrendű tartóval (egy frekvenciasávon belül azonos értékek) 256 pontossá alakítjuk, majd ezt kivonva a spektrális burkolóból kapjuk a minimális kvantálási jel/zaj viszonyt. 256 pontra számoltunk, tehát a bitkiosztás is mind a 256 frekvencia együtthatóra saját értéket fog adni. A végső bitkiosztást a megkövetelt kimeneti bitsebesség függvényében számoljuk. Az egy keret 6 audió blokkjának egyenként 5.1 csatornája közös bitkészleten osztozik. Ez nagyon rugalmas bitkiosztást tesz lehetővé. Viszont nincs mód adott keret információját szomszédos keretekre is kiterjeszteni.

### **Mantisszák kvantálása**

A számolt bitkiosztás alapján minden mantisszát önállóan kvantálunk, majd némelyeket csoportokba fogva, kódolva továbbítunk.

### **Keretképzés**

Az előzőekben leírt lépéseken keresztülmenve a bitfolyam a következőképpen alakul:



4.13. ábra. AC-3 keretszerkezete

### A kódolás hatékony eszközei

Lehetőség van a továbbított kitevők alapján számolt maszkolási görbe közvetlen módosítására kóder oldalról, explicit módosítási paraméterek továbbításával

Az alkalmazott hallásmodell 2 egyenest tartalmazó prototípus maszkolási görbét használ, mely jobb, mint az általánosan használt 1 egyenest tartalmazó modellek.

A hallásmodell paraméterei kóderi oldalról állíthatók.

Itt is lehetőség van intenzitásos sztereóra, itt azonban a résztvevő csatornák nagyfrekvenciás komponenseiből átlagolással egyetlen közös csatornát képeznek, majd az egyes csatornákat a közös csatorna alsó határfrekvenciájára sávkorlátozva továbbítják, valamint egy normál csatornához hasonlóan, a közös csatornát is kódolva továbbítják. Az egyes csatornák saját intenzitásának reprodukálására minden csatornára részsávcsopontonként koordináták adhatók meg.

Kétcsatornás üzemmód esetén Dolby Surround jelek kódolásánál 4 frekvenciasávra lebontva megadható a csatornák definíciója: B és J avagy B+J és B–J

### Továbbítható, továbbítandó segédinformációk

Közép- vagy háttércsatorna használata esetén továbbítani kell a kisebb csatornaszámon való lejátszáshoz szükséges keverési együtthatókat

Jelezni kell a bitfolyam típusát:

| bsmod<br>paraméter | csatornaszám      | szolgáltatás típusa                          |
|--------------------|-------------------|----------------------------------------------|
| 0                  | tetszőleges (5.1) | fő audió szolgáltatás: teljes szolgáltatás   |
| 1                  | tetszőleges (5.1) | fő audió szolgáltatás: csak zene és effektek |
| 2                  | tetszőleges (5.1) | járolékos szolgáltatás: képhez kapcsolt      |
| 3                  | tetszőleges (5.1) | járolékos szolgáltatás: hanghoz kapcsolt     |
| 4                  | tetszőleges (5.1) | járolékos szolgáltatás: dialógus             |
| 5                  | tetszőleges (5.1) | járolékos szolgáltatás: kommentár            |
| 6                  | tetszőleges (5.1) | járolékos szolgáltatás: vészközlemény        |
| 7                  | 1                 | járolékos szolgáltatás: alámondás            |
| 7                  | 2 – 5.1           | fő audió szolgáltatás: karaoke               |

1. táblázat. AC-3 bitfolyam típusai

Ez elsősorban a kép kísérőhangjaként szolgáló bitfolyamoknál érdekes. A képhez több AC-3 bitfolyam kapcsolható. A bennük kódolt szolgáltatások típusának megjelölésével valósítható meg pl. a többnyelvű hang továbbítása. Ez kétféleképpen történhet: egy teljes szolgáltatás (bsmod=0) mellé több dialógust (bsmod=4) teszünk, melyek csak a beszédet tartalmazzák, és a teljes szolgáltatással együtt dekódolandók, avagy több teljes szolgáltatást (bsmod=0) küldünk, mely ugyan több információ továbbításával jár, viszont könnyebb dekódolni, mivel egyszerre mindig csak egyet kell.

Egy dekódertől elvárt, hogy egy fő- és egy járulékos szolgáltatást képes legyen egyidőben dekódolni.

Meg kell adni a hanganyagban található beszéd akusztikai jelszintjét

Jelezni lehet a hanganyag nyelvét

### Továbbítható idő kód

Jelezni lehet, hogy a hanganyagot milyen típusú teremben keverték (kicsi vagy nagy), és mekkora volt az abszolút akusztikai szintje egy önálló csatornának a végső keverés során. Ezzel lehetőség nyílik a dekóderben a lejátszási abszolút hangszinttől függően a hanganyag karakterisztikáját (frekvenciamenetét) az emberi hallásgörbékhez igazítani.

dinamika-szabályzó paraméterek továbbíthatók, a gyengébb lehallgatási körülmények közötti hangzás javítására

## 4.6. Az AC-3 és az MPEG-2 többcsatornás kódolási eljárások összehasonlítása

| AC-3                                                            | Tulajdonság | MPEG-2                                                                                                      |
|-----------------------------------------------------------------|-------------|-------------------------------------------------------------------------------------------------------------|
| frekvencia- ill. időfelbontás 48 kHz-es mintavételi frekvencián |             |                                                                                                             |
| hosszú ablak: 93.8 Hz/5.3 ms<br>rövid ablak: 187.5 Hz/2.7 ms    |             | 2-es réteg: 750 Hz/24 ms<br>3-as réteg: hosszú ablak: 42.7 Hz/24 ms<br>3-as réteg: rövid ablak: 125 Hz/8 ms |
| frekvenciaminták normalizálása                                  |             |                                                                                                             |
| 24 lépcsős mintánként                                           |             | 2-es réteg: 62 lépcsős/12 mintánként                                                                        |
| bitkiosztás továbbítása                                         |             |                                                                                                             |
| nem                                                             |             | igen                                                                                                        |
| kvantált minták továbbítása                                     |             |                                                                                                             |
| folyamatosan kvázi-kódolatlanul <sup>2</sup>                    |             | 2-es réteg: folyamatosan, egyenként<br>3-as réteg: entrópia kódolással                                      |
| többcsatornás predikció                                         |             |                                                                                                             |
| nincs                                                           |             | van                                                                                                         |
| középcsatorna fantomkódolása                                    |             |                                                                                                             |

<sup>2</sup> 3 db 3 lépcsőre kvantált minta 5 biten, 3 db 5 lépcsőre kvantált minta 7 biten, 2 db 11 lépcsőre kvantált minta 7 biten összefogva, a többi kvantálási lépcsőszám esetén önállóan mintánként

| AC-3                                                                                                 | Tulajdonság | MPEG-2                                                                                                  |
|------------------------------------------------------------------------------------------------------|-------------|---------------------------------------------------------------------------------------------------------|
| előfeldolgozással megvalósítható                                                                     |             | beleépítve a kódolási eljárásba                                                                         |
| információ-szétterítés szomszédos keretekre                                                          |             |                                                                                                         |
| nincs                                                                                                |             | 2-es réteg: nincs<br>3-as réteg: megengedett                                                            |
| intenzitásos sztereó (dinamikus áthallás)                                                            |             |                                                                                                         |
| egy közös csatorna képezhető                                                                         |             | bármely csatorna használhatja egy másik csatorna mintáit a saját léptéktényezőivel                      |
| többnyelvű műsorok                                                                                   |             |                                                                                                         |
| külön bitfolyamokban                                                                                 |             | egy bitfolyamon belül                                                                                   |
| kompatibilis                                                                                         |             |                                                                                                         |
| csak kisebb csatornaszám felé                                                                        |             | MPEG-1 felé,<br>kisebb csatornaszám felé<br>(létezik MPEG-1 -el nem kompatibilis verzió is: MPEG-2 NBC) |
| mátrixolás                                                                                           |             |                                                                                                         |
| nincs                                                                                                |             | kompatibilitás miatt van,<br>kivéve a nem kompatibilis verzióban                                        |
| Dolby Surround jel figyelembevétele                                                                  |             |                                                                                                         |
| van                                                                                                  |             | van                                                                                                     |
| dinamikaszabályzás vezérlése                                                                         |             |                                                                                                         |
| szabványban definiált                                                                                |             | nem definiált                                                                                           |
| kódolási-dekódolási késleltetési idő 48 kHz-es mintavételi frekvencia esetén<br>(gyakorlati értékek) |             |                                                                                                         |
| csak fejléc-hibavédelemmel: 89 ms<br>teljes keret hibavédelmével: 101 ms                             |             | 2-es réteg: 100 ms<br>3-as réteg: 150 ms                                                                |

2. táblázat. Az AC-3 és az MPEG-2 Audio kódolások összehasonlítása

Az összehasonlítás alapján a következőket mondhatjuk: Az MPEG-2 2-es rétegű kódolás hatékonysága a legkisebb, számításigénye pedig összemérhető az AC-3-éval. A 3-as rétegű kódolás kb. hasonló minőségű tömörítést tud elérni, mint az AC-3, viszont számításigénye lényegesen nagyobb. Mindez azzal magyarázható, hogy az MPEG kódolások egymással kompatibilisek, és ez vagy a



kódolt hanganyag minőségének romlásában, vagy a számításigény jelentős növekedtében jelentkezik.

## Irodalomjegyzék

- [1] G. Davidson, L. Fielder and B. Link, Parametric Bit Allocation in a Perceptual Audio Coder. <http://www.dolby.com/tech/parametr.html>
- [2] M. Bosi and G. Davidson, High-Quality, Low-Rate Audio Transform Coding for Transmission and Multimedia Application, <http://www.dolby.com/tech/highqual.html>
- [3] E. Zwicker and H. Fastl, Psychoacoustics - Facts and Models. Springer-Verlag, Berlin, 1990.
- [4] C. Todd, G. Davidson, M. Davis, L. Fielder, B. Link, S. Vernon: AC-3: Flexible Perceptual Coding for Audio Transmission and Storage, <http://www.dolby.com/tech/ac-3flex.pdf>
- [5] International Standard ISO/IEC 11172-3: Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s. Part 3: Audio, 1992
- [6] International Standard ISO/IEC 13818-3: Generic Coding of Moving Pictures and Associated Audio Information. Part 3: Audio, November 1994
- [7] Stefan Scheller: Optimal Pulse Code Modulation (oPCM) and its application as an audio quality parameter. October 1996, <http://cips02.physik.uni-bonn.de/~scheller/aquality.html>