Open Access 2026-11-02

Siska Ádám – Szigetvári Andrea

Hangdizájn, hangszintézis és hangátalakítás

A tananyag 15 fejezetben tárgyalja a különböző hangszínterek létrehozásának lehetőségeit a legelterjedtebb szintézistechnikákon és jelfeldolgozási módszereken keresztül. Minden fejezet egy elméleti összefoglalóval indul, mely tömören és egyszerűen mutatja be az egyes témák elméleti hátterét, kiemelve a zenei szempontból lényeges mozzanatokat. Ezt Max és CSound programnyelveken készített gyakorlati példák követik, melyekkel az Olvasó maga is kipróbálhatja az egyes eljárásokat. A jobb megértést az egyes fejezeteket lezáró összegzések, valamint a témákhoz tartozó számtalan kreatív feladat segítik. A bemutatott jelenségeket, eszközöket számos kép- és hangpéldával illusztráltuk.

Kulcsszavak: zene, elektronikus zene, hangdizájn, hangszíntézis, jelfeldolgozás, spektrális analízis, multidimenzionális hangszíntér, hangszintézis technika, formáns, szűrés

BUDAPESTI KOMMUNIKÁCIÓS ÉS ÜZLETI FŐISKOLA

Typotex Kiadó

2014

Open Access 2026-11-02

COPYRIGHT: © 2014-2019, Siska Ádám, Szigetvári Andrea, Budapesti Kommunikációs és Üzleti Főiskola

Creative Commons NonCommercial-NoDerivs 3.0 (CC BY-NC-ND 3.0)

A szerző nevének feltüntetése mellett nem kereskedelmi céllal szabadon másolható, terjeszthető, megjelentethető és előadható, de nem módosítható.

Lektorálta: Kovács Balázs

ISBN 978 963 279 213 2

Készült a Typotex Kiadó gondozásában

Felelős vezető: Votisky Zsuzsa

Készült a TÁMOP-4.1.2/A/1-11/1-2011-0010 számú, „Digitális és kollaboratív művészet” című projekt keretében.



Open Access 2026-11-02

Tartalomjegyzék

Bevezetés **6**

1. A hangszín fogalma, a multidimenzionális hangszíntér **8**
1.1. Elmélet 8
1.2. Illusztráció 11
1.3. Feladatok 12

2. Hullámforma-ismétlés **15**
2.1. Elmélet 15
2.1.1. Mintavételezés 15
2.1.2. Generálás 15
2.2. Megvalósítás 17
2.3. Összegzés 18
2.4. Feladatok 20
2.5. Függelék 21

3. Additív szintézis **24**
3.1. Elmélet 24
3.2. Megvalósítás 26
3.3. Összegzés 29
3.4. Feladatok 30
3.5. Függelék 30

4. Szubtraktív szintézis **33**
4.1. Elmélet 33
4.1.1. Forrásjelek 33
4.1.2. Szűrőtípusok 34
4.2. Megvalósítás 35
4.3. Összegzés 38
4.4. Feladatok 39
4.5. Függelék 39

5. Granuláris szintézis **42**
5.1. Elmélet 42
5.2. Megvalósítás 45
5.3. Összegzés 45
5.4. Feladatok 47

© *Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola*

Open Access 2026-11-02

5.5. Függelék 48

6. Fizikai modellezés **50**
6.1. Elmélet 50
6.2. Megvalósítás 53
6.3. Összegzés 53
6.4. Feladatok 54
6.5. Függelék 55

7. Formáns szintézis **57**
7.1. Elmélet 57
7.1.1. A formánsok 57
7.1.2. A magánhangzók képzése 58
7.2. Megvalósítás 60
7.2.1. Formáns szintézis szűrőkkel 60
7.2.2. Additív formáns szintézis 61
7.3. Összegzés 65
7.4. Feladatok 66

8. Modulációs technikák I: amplitúdó- és ringmoduláció **68**
8.1. Elmélet 68
8.1.1. A folyamat 68
8.1.2. A spektrum és a hangszín 68
8.2. Megvalósítás 71
8.3. Összegzés 72
8.4. Feladatok 73
8.5. Függelék 73

9. Modulációs technikák II: frekvencia-moduláció **74**
9.1. Elmélet 74
9.1.1. A folyamat 74
9.1.2. A spektrum és a hangszín 76
9.1.3. Álfrekvenciák keletkezése 79
9.2. Megvalósítás 80
9.3. Összegzés 82
9.4. Feladatok 82
9.5. Függelék 83

10. Spektrumanalízis **85**
10.1. Elmélet 85
10.2. Megvalósítás 89
10.3. Összegzés 90
10.4. Feladatok 90

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

11. Reszintézis és alkalmazásai: transzpozíció, nyújtás, szűrés, keresztszintézis 92
11.1. Elmélet . . . 92
11.1.1. Reszintézis . . . 92
11.1.2. Nyújtás . . . 93
11.1.3. Transzponálás . . . 94
11.1.4. Spektrális szűrés . . . 95
11.1.5. Keresztszintézis . . . 95
11.2. Megvalósítás . . . 97
11.3. Összegzés . . . 98
11.4. Feladatok . . . 99

12. Jelfeldolgozás I: késleltetésen alapuló eszközök 101
12.1. Elmélet . . . 101
12.1.1. Szűrés vagy késleltetés? . . . 101
12.1.2. Flanger, fázistoló . . . 102
12.1.3. Elhangolás, vibrato . . . 103
12.1.4. Kórus . . . 104
12.2. Megvalósítás . . . 105
12.3. Összegzés . . . 108
12.4. Feladatok . . . 108

13. Jelfeldolgozás II: torzítók 110
13.1. Elmélet . . . 110
13.1.1. Jelalak-torzítás . . . 110
13.1.2. Kvantálási torzítás . . . 111
13.2. Megvalósítás . . . 113
13.3. Összegzés . . . 114
13.4. Feladatok . . . 114

14. Jelfeldolgozás III: kompresszálás, zengetés 116
14.1. Elmélet . . . 116
14.1.1. Kompresszálás . . . 116
14.1.2. Zengetés . . . 120
14.2. Megvalósítás . . . 122
14.3. Összegzés . . . 124
14.4. Feladatok . . . 124
14.5. Függelék . . . 125

15. Multidimenzionális hangszínvektorok 128
15.1. Elmélet . . . 128
15.1.1. Hangszínvektorok . . . 128
15.1.2. Pontok, trajektóriák, kombinációk . . . 129
15.2. Megvalósítás . . . 130
15.3. Összegzés . . . 131
15.4. Feladatok . . . 132

Bibliográ a 133

© *Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola*

Open Access 2026-11-02

Bevezetés

Ebben a tananyagban azt mutatjuk be, hogy hogyan lehet különböző (konstans és időben változó) hangszíneket létrehozni. Ehhez különböző *szintézistechnikákat*, illetve *jelfeldolgozási módszereket* ismertetünk.

Szintézistechikának hívunk minden olyan eljárást, amely során a „semmiből" hozunk létre hangokat. Egy szintézistechnikát egyrészt a *szintéziselv*, másrészt a *szintézis paraméterei* határoznak meg. Szintéziselv alatt többnyire azt az algoritmust értjük, amely az adott hangot létrehozza; a paraméterteret az adott algoritmusnak a változói alkotják. Az egyes szintézismódszereket az általuk létrehozható hangszíntér jellemzői, a paramétereik mennyisége és „bonyolultsága,[1]" illetve az algoritmus erőforrásigénye alapján osztályozzuk. Jelenleg nem ismerünk olyan szintézistechnikát, amely képes lenne tetszőleges hangszínt úgy előállítani, hogy eközben a paraméterei könnyen átláthatóak maradjanak. Emiatt a helyes szintézismódszer megválasztásának döntő jelentősége van egy-egy konkrét hangszín megalkotásakor.

A jelfeldolgozási eljárások („*effektek*") ezzel szemben mindig létező hangokból indulnak ki; hatásukat e hangok manipulálása által érik el. Emiatt az általuk elért hangszín alapvetően függ a beérkező hangszíntől. Az egyes módszerek között aszerint teszünk különbséget, hogy a beérkező hangszíneket milyen módon (milyen hangszíntér-dimenziók mentén) módosítják.

A szintézistechnikák és a jelfeldolgozási módszerek között nincs éles határ. Bizonyos algoritmusok (pl. a granuláris szintézis vagy a frekvencia-moduláció, lásd az 5. vagy a 9. fejezetet) a paramétereik függvényében egyaránt viselkedhetnek szintézismódszerként vagy jelfeldolgozási eljárásként.

Az 1. fejezet a hang alapvető jellemzőit tárgyalja, elsősorban a hangszín, illetve a hangszínt leíró legfontosabb paraméterek szempontjából.

A 2–7. fejezetek mutatják be a legelterjedtebb szintézistechnikákat. Ezek mindegyike ún. *konvencionális* szintézistechnika, melyek valamilyen (pszicho)akusztikus hangzási modellből indulnak ki.[2]

A 8–11. fejezetek olyan technikákat ismertetnek, melyek túlmutatnak az egyszerű hangszintézisen, illetve jelfeldolgozáson. A 8–9. fejezetek az amplitúdó- és frekvenciamodulációt mutatják be, melyek (ahogy már említettük) egyaránt tekinthetők szintézistechnikának és jelfeldolgozási módszernek. A 10–11. fejezetek olyan módszereket tárgyalnak, melyek a hang analízisén, illetve az ebből nyert (és adott esetben módosított) adatsorokat felhasználó hang-rekonstrukción alapulnak.

A 12–14. fejezetek ismertetik a leggyakoribb jelfeldolgozási módszereket. Ezen belül a 12–13. fejezetek a szintézis (vagy hangfelvétel) során kapott hangok alapvető manipulációit, míg a 14. fejezet a már manipulált hangok végső keverését taglalják.

[1] Ez alatt azt értjük, hogy a paraméterek mennyire „szemléletesen" befolyásolják a hangszínt.

[2] A nemkonvencionális technikák ezzel szemben a hangkeltés szempontjából absztraktnak számító folyamatokon – pl. tőzsdei árfolyam-ingadozások közvetlen szonifikációján – alapulnak.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

A 15. fejezet a hangszíntér általános leírhatóságának és kezelésének a kérdéseit vizsgálja.

Mindegyik fejezet egy-egy elméleti bevezetővel indul, amely a bemutatott eljárások célját, alapvető viselkedését és lehetséges paramétereit (illetve azok jelentését) tárgyalja. Ezt követően különböző példaprogramokat láthatunk, melyek a mondottakat illusztrálják. Ezek egy része CSound nyelven írt forráskód, a többség azonban önállóan futtatható szoftver. A példaprogramok elsődleges célja az egyes jelenségek bemutatása, így a hangsúly az egyszerű kezelhetőségre és az áttetszőségre került; e programok természetesen nem tudják és nem is igyekeznek felvenni a versenyt a kereskedelmi forgalomban kapható professzionális eszközökkel. Sőt, számos esetben olyan lehetőségeket és beállításokat is tartalmaznak, melyeket egy professzionális szoftver eleve kizár. Célunk az volt, hogy az extrém paramétereken és beállításokon keresztül jobban rávilágíthassunk egy-egy módszer természetére.

A bemutatott technikákat – ahol ennek létjogosultsága van – három szempontcsoport szerint értékeljük minden egyes fejezet végén. Összefoglaljuk, hogy az adott eljárással milyen hangszínek, illetve hangszínmódosítások érhetőek el; ismertetjük a szükséges paraméterek számát és jellegét, és röviden értékeljük, hogy ezek milyen kapcsolatban állnak a létrehozható hangszíntérrel; végül említést teszünk az egyes módszerek erőforrás-igényéről. A fejezeteket az elmondottak jobb megértését és gyakorlását szolgáló feladatokkal zárjuk.

A bemutatott CSound-példákat az egyes fejezetek végén, külön Függelékben közöljük, míg a Max példaprogramokat képernyőfotók segítségével illesztettük a tananyagba. Az elmondottakat, ahol szükségesnek láttuk, hangpéldákkal illusztráltuk.

© *Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola*

Open Access 2026-11-02

1. fejezet

A hangszín fogalma, a multidimenzionális hangszíntér

1.1. Elmélet

Fizikai értelemben a hang az anyag rezgéseit jelenti. Egy rezgést *szabályosnak*, vagy *periodikusnak* hívunk, ha az szabályos időközönként megismétli önmagát. A legrövidebb időtartam, ami alatt ez az ismétlődés bekövetkezik, a rezgés *periódusideje*; az így ismétlődő rezgésmintázat pedig a *periódus*. A periódusidő helyett gyakran annak inverzét, a *rezgésszámot* vagy *frekvenciát* használjuk, ami tehát azt mutatja meg, hogy egységnyi idő alatt hány periódus hangzik el.[1] A rezgés erősségét az egyensúlyi helyzettől mért legnagyobb kitérés, az *amplitúdó* jellemzi.

Az emberi észlelés az egyszerűségre törekszik, ami abban nyilvánul meg, hogy az érzékszerveink által felfogott jelekben mintázatokat keresünk, a valóság pedig e mintázatokon keresztül tudatosul bennünk. Mivel egy szabályos rezgés két adatra redukálható (a frekvenciára és az amplitúdóra), ezért e rezgések központi szerepet játszanak a zenei észlelés során. Az ilyen jeleket tudatunk egy konkrét *hangmagassággal* és *hangerővel* ruházza fel, melyek a frekvenciától és az amplitúdótól függnek: minél mélyebb egy hang, annál kisebb a frekvenciája, illetve minél nagyobb az amplitúdója, annál hangosabb.

A hangmagasságon és a hangerőn kívül természetesen maga a periódus, illetve annak belső szerkezete az, ami a szabályos rezgést leginkább jellemzi. Ez nem írható le egy egyszerű paraméterrel, hiszen itt egy (adott esetben igen komplex) struktúráról beszélünk. Ennek megfelelően a periódus egy jóval összetettebb fogalommal társul a tudatunkban: ez maga a *hangszín*.

Különböző (szabályos) hangok keveredésekor egy összetett, de szintén periodikus (vagy legalábbis majdnem-periodikus) rezgést kapunk. Ezt a jelet tudatunk három módon dolgozhatja fel. Egyrészt dönthet úgy, hogy egyetlen hangként fogja fel, saját hangmagassággal, hangerővel és hangszínnel. Másrészt azt is választhatja, hogy több, önálló hangot észlel a hallott jelben. Előbbi esetben *összeolvadásról*, utóbbi esetben *szétválásról* beszélünk. Hogy agyunk pontosan milyen kritériumok alapján dönt a két megoldás között (vagyis, hogy melyik megoldást tekinti egyszerűbb reprezentációnak), jelenleg is kutatás tárgya, habár vannak bizonyos eredményeink. Ezekről bővebben a 3. fejezetben ejtünk szót.

[1] Mértékegysége a *hertz* (Hz), mely az egy másodperc alatti rezgések számát adja meg.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

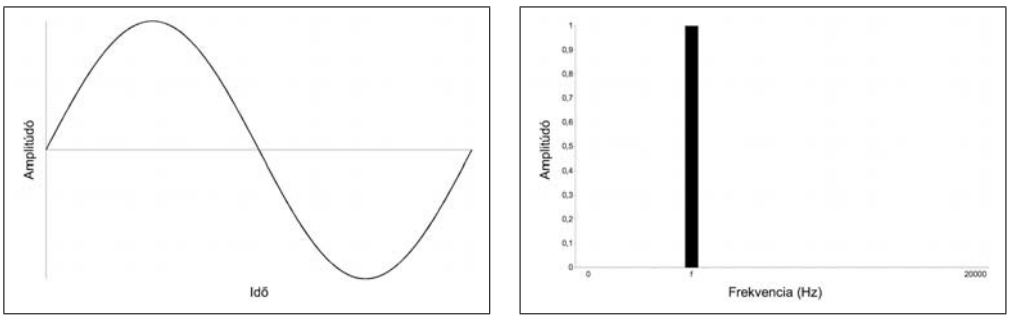
Amennyiben egy rezgés egyáltalán nem mutat semmilyen periodicitást, vagy túl sok egyenrangú komponensből épül fel, tudatunk nem rendel hozzá konkrét hangmagasságo(ka)t (ez a harmadik eset, az előző bekezdés gondolatmenetét folytatva); az ilyen jelet zajként értelmezzük. Ám agyunk ilyenkor is hasonló mintázatokat keres, mint amikor a „szabályos” hangok periódusát elemzi (lásd alább), ezáltal képesek vagyunk különbséget tenni a különböző zajok között is. Ezt illusztrálja a következő három (konkrét hangmagassággal nem rendelkező, zajos) hangpélda:

1.1 : Három különböző dob egymás után.

1.2 : Fehér zaj (ennek definícióját lásd a 4. fejezetben).

1.3 : Recsegő, zajos hangok.

Matematikai okokból a különböző szerkezetű periódusok között a szinuszos rezgések (lásd az 1.1. ábrát) kiemelt helyet foglalnak el, ugyanis tetszőleges rezgés felbontható ezek összegére. Ezt a felbontást a hang *spektrumának* hívjuk; magát a hangszínt e spektrummal azonosítjuk. Az 1.2. ábra egy komplex hang spektrumát illusztrálja.



1.1. ábra. Szinuszos rezgés. Az egyensúlyi helyzettől való kitérés az idő függvényében (bal), illetve a szinuszos rezgés spektruma (jobb). A szinuszos rezgést az 1.4. hangpélda mutatja be.

Habár a spektrum (matematikai értelemben) a lehető legprecízebben mutatja be a hangszínt, bonyolultsága miatt a zenei gyakorlatban nem ezt a leírást használjuk, hanem ebből származtatható, egyszerűbb paramétereket. Ez összhangban áll agyunk működésével is: az egyszerű leírásra kihegyezett tudatunk a spektrumban mintázatokat keres, és ezek alapján kategorizálja a hangszíneket. E mintázatokat, pontosabban az ezeket leíró jelzőket nevezzük a hangszín *dimenzióinak*.

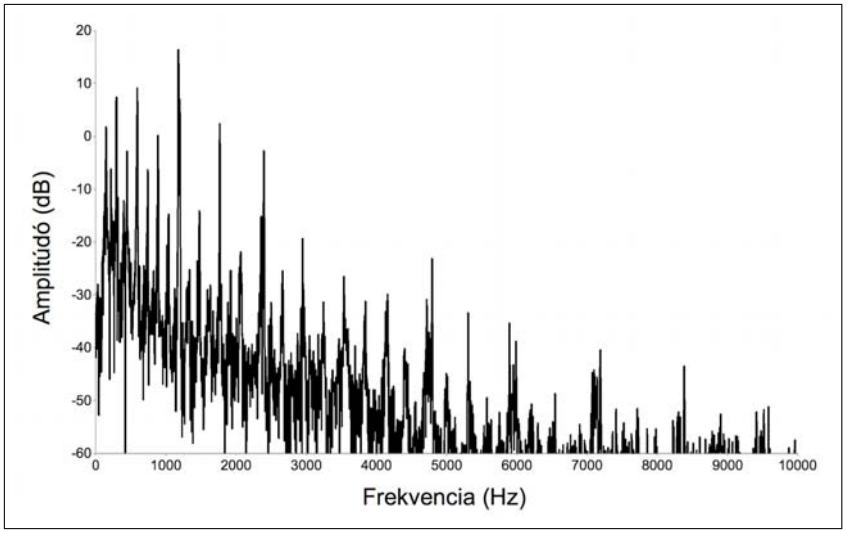
A hangszíntér dimenzióinak meghatározása – az egyes dimenziók azonosítása, illetve ezek származtatása a spektrumból – jelenleg is aktív kutatási terület. Olyan paraméterekre kell gondolnunk, mint például a *zajosság*, *fényesség*, *érdesség*, *fémesség*, *élesség* stb. Ezek egy része közvetlenül kiszámítható a spektrum ismeretében[2] – természetesen csak közelítő jelleggel, hiszen (a hangmagasságokkal vagy a zenei időtartamokkal ellentétben) ezek a jellemzők erősen szubjektívek.

Az eddig elmondottak során hallgatólagosan feltételeztük, hogy az általunk vizsgált hang időben állandó: a periódus változatlan formában, végtelenül sokszor ismétlődik benne. Mivel a valóságban ez nincs így, általában rövid (legfeljebb néhány századmásodperces) sza-

[2] A spektrum meghatározásának módszereit a 10. fejezetben ismertetjük.

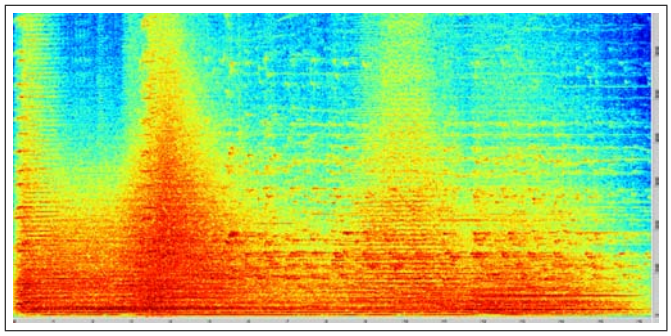
© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



1.2. ábra. Egy zenekari hang spektruma (a hangot az 1.5. hangpélda tartalmazza). Az egyes szinuszos rezgések amplitúdóját a függőleges, frekvenciáját a vízszintes tengely mutatja.

kaszokra bontjuk a vizsgált jelet, és az így kapott csonkokra úgy tekintünk, mintha egy idealizált, „végtelen” hang részletei lennének. Így szakaszonként meghatározhatóvá válik a „pillanatnyi spektrum”. A jel ilyen módon való ábrázolását *szonogramnak* hívjuk, melyet az 1.3. ábra szemléltet.



1.3. ábra. Egy zenekari részletről készített szonogram (a hangot az 1.6. hangpélda tartalmazza). A függőleges tengely a frekvenciát (hertz), a vízszintes az időt (másodperc) mutatja. A pillanatnyi spektrumot (az adott frekvenciák pillanatnyi erősségét a hangban) az alkalmazott kék-piros színskála érzékelteti.

A szonogram azért is szerencsés választás a hang általános ábrázolására, mert jól képezi le azt a folyamatot, ami tudatunkban is végbemegy a fülünk által észlelt hangjelek értelmezésekor. Habár nem tudjuk pontosan, hogy ez az értelmezés hogyan zajlik, úgy tűnik, hogy a „szakaszokra bontás” fent ismertetett (a 10. fejezetben részletesebben kifejtett) módszere nem állhat messze természetes hangfelfogásunktól. Ezt illusztrálják az 1.7. és az 1.8. hangpéldák, melyek ugyanazt a zongorahangszínt tartalmazzák eltérő amplitúdó-burkológörbékkel. Habár a két hang ugyanazokból a spektrális komponensekből áll, az 1.7. hangpélda kezdetben hirtelen változó burkológörbéje zajossá teszi a hangindítást; a két hangszín emiatt lényegesen eltér. A jelenséget csak úgy tudjuk megmagyarázni, ha

Open Access 2026-11-02

feltételezzük, hogy tudatunk is egy, a szonogram-ábrázoláshoz hasonló „pillanatnyi spektrumot” érzékel, hiszen ez teszi lehetővé, hogy az 1.7. példában a hangindításkor fellépő tranzienseket valóban zajszerűnek észleljük.

Felmerül a kérdés: meddig tart egy pillanat? Vagyis: mi az az időtartam, ami alatt a fülünk a jelet egyetlen hangnak fogja fel, és ami felett képesek vagyunk időbeli folyamatokat észlelni? A kérdés azért bonyolult, mert maga a hangmagasság is az idő (konkrétan: a periódusidő) következménye, így nem beszélhetünk külön időről és hangmagasságról; a kettő „egylényegű”. Ahhoz, hogy tudatunk egyértelműen megállapíthassa egy rezgés hangmagasságát, néhány periódusnyit hallanunk kell a hangból. Ebből következik, hogy azokat a zenei folyamatokat, amelyek nagyságrendekkel lassabban mennek végbe, mint a hangok jellemző periódusidejei,[3] a tudatunk valóban időbeli folyamatokként fogja értelmezni (ilyenek pl. a zenei artikulációhoz kapcsolódó folyamatok, mint a ritenuto vagy a crescendo). Azonban az ennél gyorsabb folyamatok (pl. a zongora hangindításakor tapasztalt rövid tranziensek, lásd az 1.7. hangpéldát) a tudatunkban nem folyamatként, hanem hangszínként jelennek meg.

1.2. Illusztráció

A fejezet főbb megállapításait a `HApp_01_01` `(OSX)` `(WIN)` interaktív példa foglalja össze (lásd az 1.4–1.6. ábrákat). Ez egy szonogramot és egy spektrogramot tartalmaz (utóbbi a hang pillanatnyi spektrumát mutatja). A képernyő alján levő két számdobozzal határozható meg a kijelzendő spektrum legmélyebb és legmagasabb frekvenciája. A szonogram fölötti „*Megállítás*” kapcsoló bekapcsolásával a szonogram megállítható, egyébként az iménti gombtól jobbra található „*Intervallum*” határozza meg a kirajzolás sebességét. A bal oldalon levő felső vezérlő különböző hangforrásokat (mikrofon/vonalbemenet, felvett file, szinuszjel, fehér zaj és rövid impulzus) tartalmaz, míg az alsó vezérlő a kimeneti hangerőt szabályozza. A programot a „*Be/Ki*” kapcsolókkal lehet be- és kikapcsolni.

A kijelző tetején található fülek közül a bal oldali („*Alapok*”) lehetővé teszi, hogy különböző hangforrások spektrumát és szonogramját megjeleníthessük.

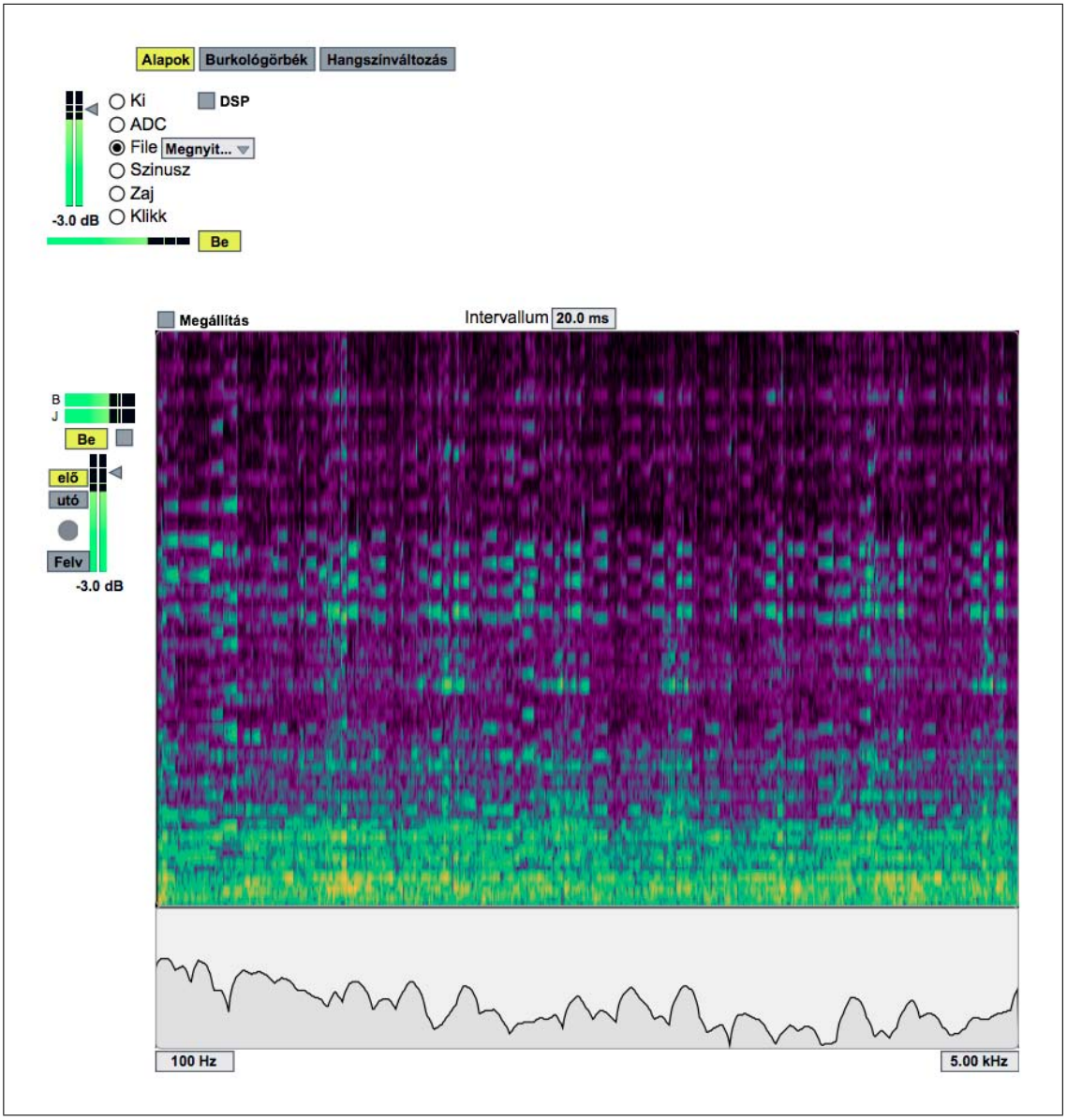
A „*Burkológörbék*” fül egy burkológörbe-szerkesztőt jelenít meg. Amennyiben a szerkesztő alatti „*Folytonos*” üzemmódból „*Szakaszos*” módba váltunk, a hang csak akkor fog szólni, ha a burkológörbe feletti „*Indít*” gombot megnyomjuk. A burkológörbe hosszát a görbe feletti számdobozzal, míg alakját az egérrel, illetve a burkológörbe melletti kapcsolók közötti váltással szerkeszthetjük. Ezen a fülön magunk is megtapasztalhatjuk, hogy a burkológörbe megváltoztatása ugyanazt az eredeti hangot teljesen különböző hangszínekkel ruházza fel, amint azt az 1.7. és az 1.8. hangpéldákban is hallhattuk.

A „*Hangszínváltozás*” fül egy alternatív hangforrással dolgozik. Itt a hangunk frekvenciáját határozhatjuk meg, valamint a periódus szerkezetét, mely a két megjelenő csúszka közül az alsóval szabályozható. Az „1”-es állapot egy tökéletes szinuszos, míg a „2”-es állapot egy komplex periódust tartalmaz. E fül segítségével kipróbálhatjuk, hogy a periódus-szerkezet változtatásával hogyan változik meg a hang spektruma (illetve hangszíne), miközben hangmagassága változatlan marad. Emellett megfigyelhetjük, hogy a burkológörbék változtatása hogyan befolyásolja ezt a teljesen egyszerű hangszínt.

[3] 20 Hz esetén a periódusidő 50 ms, 20 kHz esetén 0,05 ms. Az egyvonalas „A” (440 Hz) periódusideje közelítőleg 2,27 ms.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



1.4. ábra. `HApp_01_01` `(OSX)` `(WIN)`*: Alapok. A nagy kijelző az éppen hallott hang szonogramját, az alatta levő pedig a pillanatnyi spektrumát mutatja. A szonogram, illetve a spektrum teteje és alja szabadon állítható.*

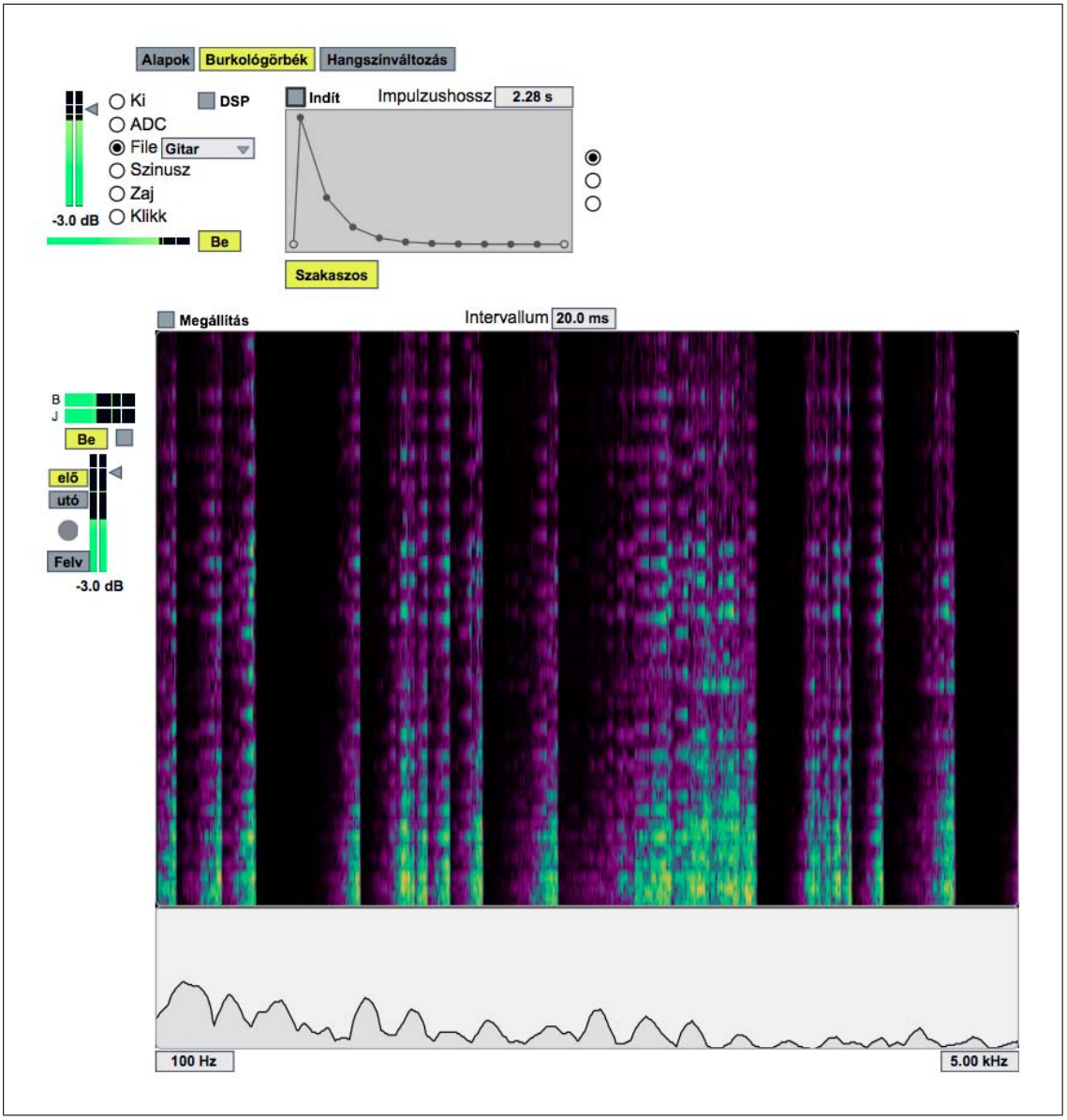
1.3. Feladatok

1. *Hallgatási gyakorlat.* Egy héten keresztül hallgass meg minél többféle hangot. Határozz meg legalább hét különböző jelzőt (hangszíntér-dimenziót), amelyek mentén a hallott hangok csoportosíthatóak! Hozz létre, vagy vegyél fel legalább tíz olyan hangot, melyek különböző kategóriákba sorolódnak az általad választott hangszíntér-dimenziók mentén!

2. A `HApp_01_01` `(OSX)` `(WIN)` példaprogram segítségével figyeld meg különböző hangok szonogramját. Állapítsd meg a szonogramképek alapján, hogy melyik hang zajos és melyik harmonikus (ha szükséges, használd a *Megállítás* opciót)! Figyeld meg,

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

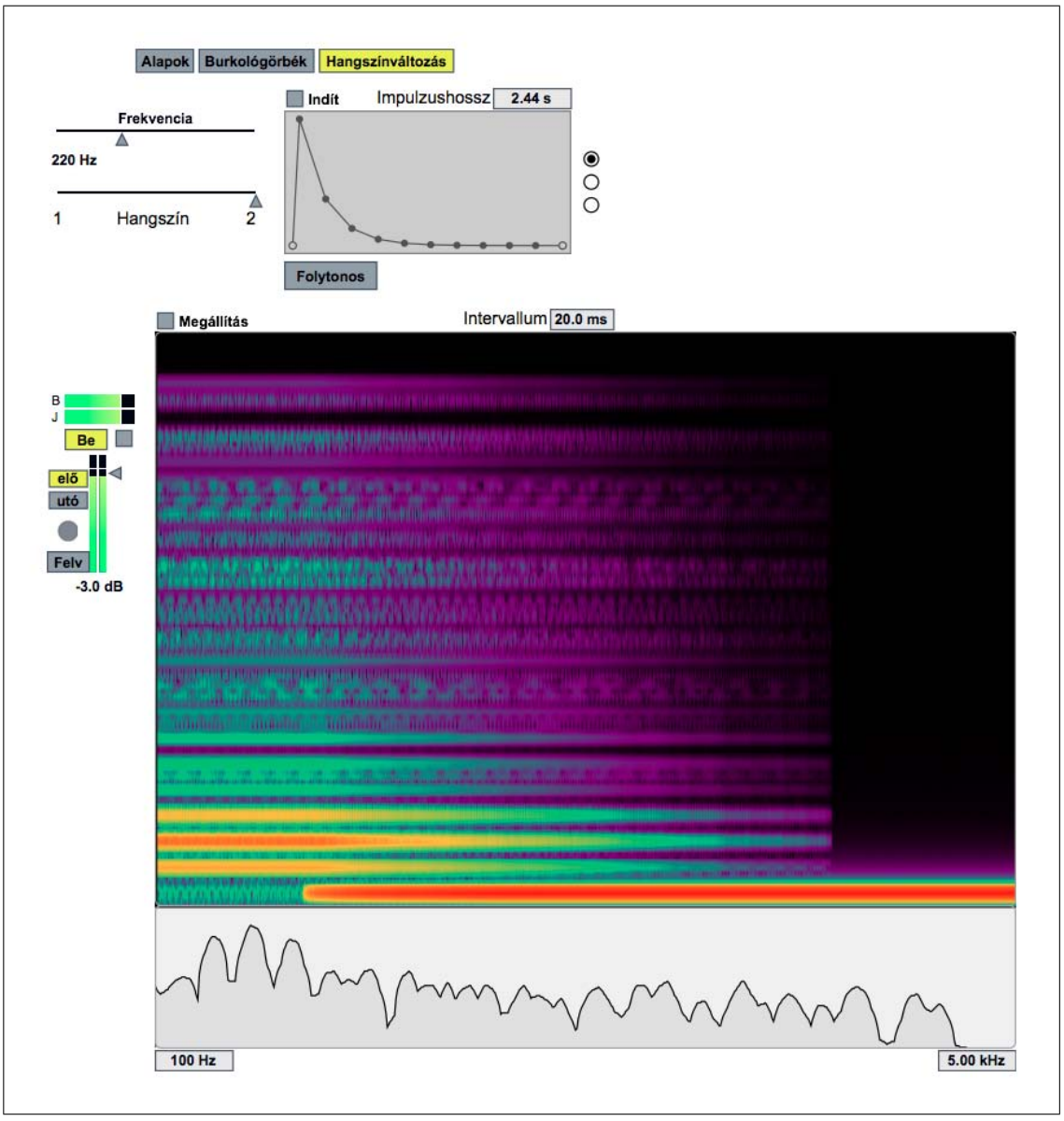


1.5. ábra. `HApp_01_01` `(OSX)` `(WIN)`*: Burkológörbék. Szakaszos üzemmódban a hang amplitúdóját tudjuk a burkológörbével szabályozni.*

hogy a szonogramban megjelenő önálló szinuszos komponensek (melyeket egy-egy vízszintes, világos csík jelöl a szonogramban) milyen esetekben olvadnak össze egyetlen hangszínné, és mikor nem!

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



1.6. ábra. `HApp_01_01` `(OSX)` `(WIN)`*: Hangszínváltozás. A két hangszín („1”-es és „2”-es) egy szinuszos hang, illetve egy fúvós hangszer hangszíne. A csúszkával tudjuk szabályozni az átmenetet e két hangszín között.*

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

2. fejezet

Hullámforma-ismétlés

Az 1. fejezetben láttuk, hogy hangmagasság-érzetünk a hang periodikus mivoltával függ össze, míg egy (harmonikus) hang hangszínét e periódus szerkezete határozza meg. A *hullámforma-ismétlés* közvetlenül aknázza ki ezt a két jelenséget.

2.1. Elmélet

Az eljárás során először meghatározzuk a létrehozandó hang egy periódusát (egy *hullámformát*), majd az így létrehozott struktúrát ismételjük. Az így kapott hang frekvenciáját az ismétlések gyakorisága határozza meg (vagyis az, hogy az adott hullámformát egységnyi idő alatt hányszor játsszuk le), míg a hangszínt a hullámforma alakja.

2.1.1. Mintavételezés

A legkézenfekvőbb eljárás, ha egy felvett hangból, mint amit a 2.1. ábra mutat, kivágunk egy teljes periódust, és ebből készítjük el a hullámformát. Az így kapott hang az eredetihez hasonló, ámde kimerevített, élettelen hangszínű lesz. A módszer segítségével időben változó hangzás különböző állapotait (kitartott harmonikus, átmeneti zajos szakaszok, stb.) lehet modellezni, és belőlük hangszínteret építeni. A periódus izolálásának „hibája”, azaz ha nem pontosan zéruskereszteződéseknél vágjuk ki a periódust, kisebb-nagyobb mértékű zaj hozzáadását okozza. A vágás helye és a zéruskereszteződés közötti távolság változtatásával folyamatos átmeneteket lehet létrehozni zajos és harmonikus hangok között. Az átmenetek nem lineárisak, kísérleti úton lehet tapasztalatot szeretni pontos működésükről.

2.1.2. Generálás

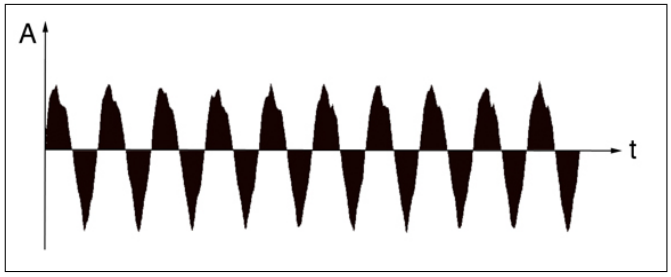
Hogyan alkothatunk meg egy hullámformát? A megoldás kulcsát az 1. fejezetben már leírtuk: „tetszőleges rezgés felbontható szinuszhullámok összegére”. Az állítást megfordítva adódik, hogy a hullámformát különböző periódusidejű és amplitúdójú szinuszhullámok összeadásával hozzuk létre.

Két dolgot érdemes végiggondolni:

- Nem lenne kötelező ragaszkodnunk a szinuszhullámokhoz, bármilyen más (periodikus) hullám is szolgálhatna alapul. Ez olyannyira igaz, hogy sokszor élünk is ezzel a

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



2.1. ábra. Fuvola hangjának részlete. A hang periodikus jellege jól látszik. Figyeljük meg, hogy az egyes periódusok nem teljesen egyformák! Ezen apró uktuációk miatt halljuk a hangot természetesnek, organikusnak.

lehetőséggel. Ám akármiből is építkezünk, a végeredményt össze lehet rakni szinuszokból is (az idézett állítás következtében). Ehhez az általunk használt alapformákat kell külön-külön felbontani szinuszos összetevőkre.

- A végeredmény csak akkor lehet periodikus, ha minden egyes összetevője is az. Ezért csak olyan szinuszt használhatunk, aminek a periódusával pontosan ki tudjuk tölteni a létrehozandó hullámformát. Ebből következik, hogy a felhasznált szinuszok frekvenciája csak a teljes hullámforma frekvenciájának egész számú többszörösei lehetnek. Az így kapott frekvenciákat az alapfrekvencia *részhangjainak*, illetve *felhangjainak*[1] hívjuk.

Az eddigiek azt a benyomást kelthették, hogy csak „szép”, ránézésre is szabályos hullámformákat lehet így létrehozni. Ez természetesen nem igaz. A felhangok megfelelő kombinációival teljesen szabálytalan formák is létrehozhatóak, extrém esetben akár törések vagy teljesen egyenes szakaszok is kialakíthatóak a hullámformában (ezek kattanó, zajos hangzásokkal gazdagítják a hangszínt). Mivel a törések és az egyenes szakaszok szinuszos kikeverése nagyon bonyolult számolást igényel, egyszerűbb kézzel „berajzolni”, vagy matematikai formulákkal generálni a hullámforma ilyen részleteit. Ezért a hullámforma generálásakor érdemes a három módszert (szinuszok összeadása, rajzolás, illetve matematikai úton való definiálás) keverni.

„Matematikai formula” alatt általában egy függvényt értünk, amely a kitérés értékét adja meg az idő függvényében. Mivel a hullámformákat ismételjük, ezért elég, ha a függvény értelmezési tartománya egyetlen hullámformára korlátozódik (például a $t = 0$ időpillanatban kezdődőre). A jelgenerátorok által létrehozott leggyakoribb hullámformák (a szinusz mellett, természetesen) a következők:

Háromszögjel:
$$y(t) = \begin{cases} 4t, & 0 \le t < \frac{1}{4} \\ 4\left(\frac{1}{2} - t\right), & \frac{1}{4} \le t < \frac{3}{4} \\ 4(t-1), & \frac{3}{4} \le t < 1 \end{cases}$$

Fűrészfogjel:
$$y(t) = \begin{cases} 2t, & 0 \le t < \frac{1}{2} \\ 2(t-1), & \frac{1}{2} \le t < 1 \end{cases}$$

[1] A kettő közti különbség, hogy az alaphang egyben részhang is, de nem tekintjük felhangnak.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

Négyszögjel: $$y(t) = \begin{cases} 1, & 0 \le t < \frac{1}{2} \\ -1, & \frac{1}{2} \le t < 1 \end{cases}$$

Impulzussorozat:[2] $$y_\delta(t) = \begin{cases} 1, & 0 \le t < \delta \\ 0, & \delta \le t < 1 \end{cases} \qquad 0 < \delta < 1$$

A fenti kifejezések mindegyike 1–periódusú függvényeket definiál, vagyis olyanokat, amelyek periódusideje egységnyi. Az egyes hullámformák egy-egy periódusát és spektrumát a 2.2. ábra ismerteti.

2.2. Megvalósítás

A módszert a `HApp_02_01` `(OSX)` `(WIN)` példa mutatja be. A felső legördülő menüből választhatunk a módszerek közül. A *Szinusz*, *Háromszög* és *Fűrészfog* opciók egy-egy, algoritmikusan definiált alap-hullámformát tartalmaznak (2.3. ábra, 2.1. hangpélda). E formulákat (a szinusz kivételével) a bal oldali panel ismerteti. A *Rajz* opció választásakor a bal oldali panelben, az egér segítségével magunk adhatjuk meg a pontos hullámformát (2.4. ábra, 2.5. hangpélda).

A *Szinuszösszeadás* során a bal oldali panelen állíthatjuk be az egyes komponensek frekvencia- és amplitúdó-arányait (2.5. ábra, 2.6. hangpélda). Úgy tűnhet, hogy ellentmondásba keveredtünk az előző szakaszban leírtakkal, hiszen az ott elmondottakkal szemben a program megengedi a nem-egész frekvenciaarányokat is. Ám megfigyelhetjük, hogy nem-egész arányok megadásakor a hullámtábla végén éles levágás keletkezik. Ahogy utaltunk rá, a „tiszta” felhangok megfelelő arányú keveredésével ezek a törések modellezhetők lennének; vagyis arról van szó, hogy – ahelyett, hogy több száz szinuszból kevernénk ki ezeket a levágásokat – egyszerűsítettünk a tervezés során. Ám fontos megértenünk, hogy ilyenkor a kapott hang spektrumában nem az eredetileg megadott szinuszok lesznek jelen, hanem az így létrehozott torzításokat leíró, bonyolultabb struktúrák.

A mintavételezéses eljárást a *Hangfájl* opció mutatja be (2.6. ábra, 2.7. hangpélda). A bal oldali panelben kiválaszthatjuk az eredeti hangokat tartalmazó file-okat, majd az egérrel kijelölhetjük a mintavételezni kívánt területet. A „*Kivágás*” gomb megnyomásakor a hullámforma a kijelölt területre szűkül. Ezt a „*Normalizálás*” gombbal normalizálhatjuk.

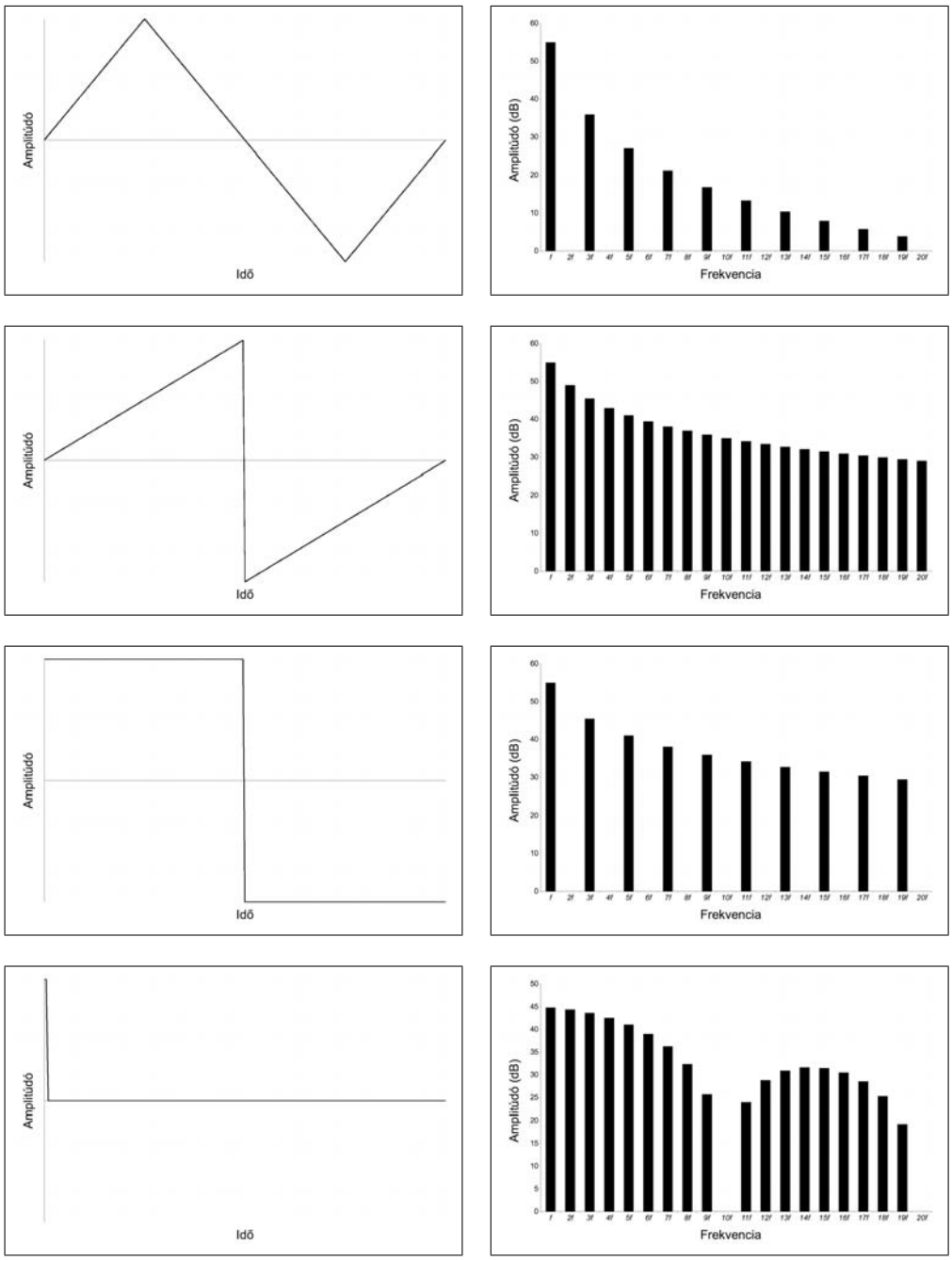
A lejátszási frekvenciát a hangerőszabályozó feletti számdobozban állíthatjuk be (hertzben kifejezve). A jobb oldali panel a kapott hangok szonogramját rajzolja ki.

A `hd_02-01_csound.csd` CSound program (lásd a 2.1. forráskódot) különböző hullámformákból álló rövid hangokat hoz létre, melynek eredményét a 2.8. hangpélda tartalmazza.

[2] Az impulzussorozat tehát nem egyetlen hullámforma, hanem egy hullámforma-család. Figyeljük meg, hogy – a többi alap-hullámformával ellentétben – az impulzussorozat időátlaga nem nulla. Ettől függetlenül, mivel a gyakorlati esetekben $\delta \ll 1$, a jelátlag nullától való eltérése elhanyagolható. További érdekesség, hogy a $\delta = \frac{1}{2}$ esetben egy (nem nulla időátlagú) négyszögjelet kapunk.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



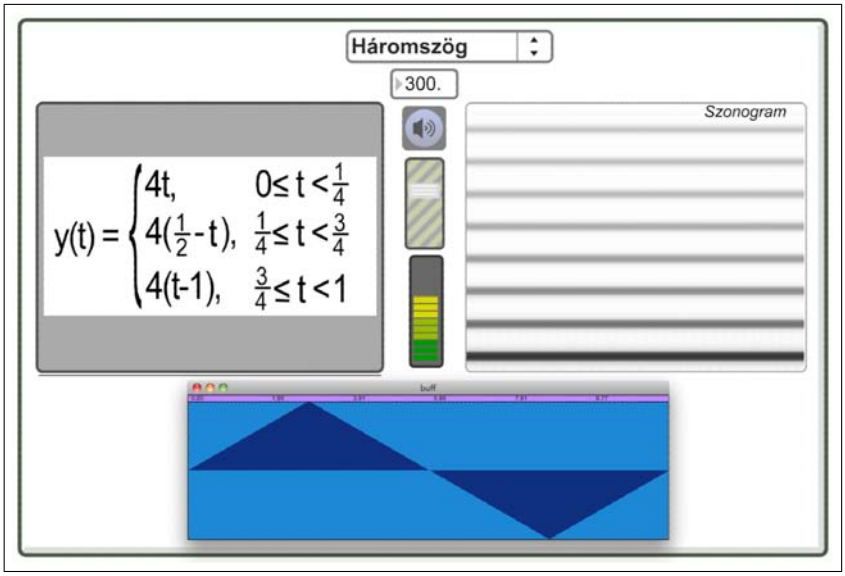
2.2. ábra. Különböző forrásjelek hullámformái (bal) és spektrumai (jobb). Fentről lefelé: háromszögjel (2.1. hangpélda), fűrészfogjel (2.2. hangpélda), négyszögjel (2.3. hangpélda) és impulzussorozat (2.4. hangpélda). A spektrumokat $L_0 = 55$ dB-re normáltuk. Az impulzussorozat hullámformája és spektruma δ-függő; ezen az ábrán $\delta = 0{,}1$.

2.3. Összegzés

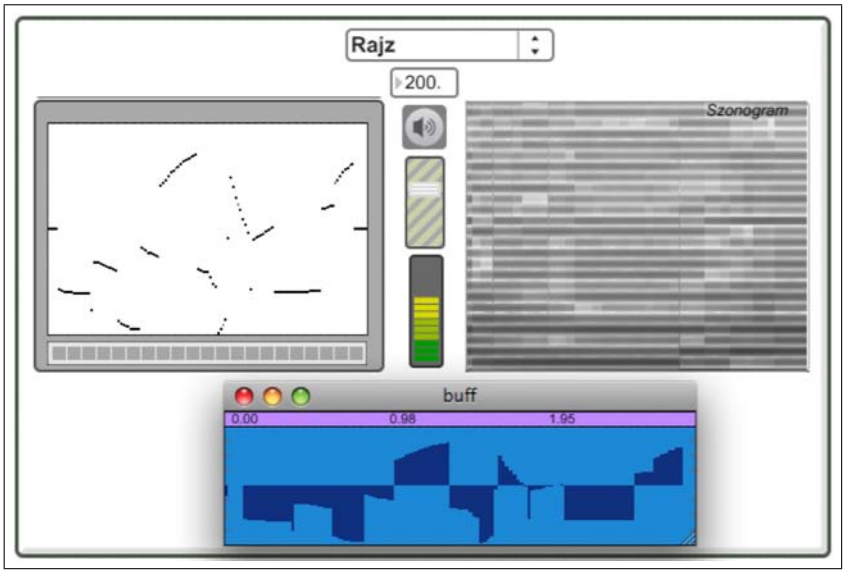
Hangszín. Tetszőleges statikus hangszín létrehozható a kiszemelt hang részletének megfelelő mintavételezésével. Mivel az eltárolt hullámformát általában nehéz élőben változtatni, ezért a végeredmény gyakran merev, élettelen. Egyes rendszerekben egy hang

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



2.3. ábra. Hullámforma-ismétlés háromszögjelből felépítve. Hasonlóan működik a program a Szinusz *és a* Fűrészfog *opciók esetén is. A bal oldalon megjelenő formula azt de niálja, hogy (egyetlen, egységnyi időtartamú periódus esetén) az amplitúdó hogyan függ az időtől.*



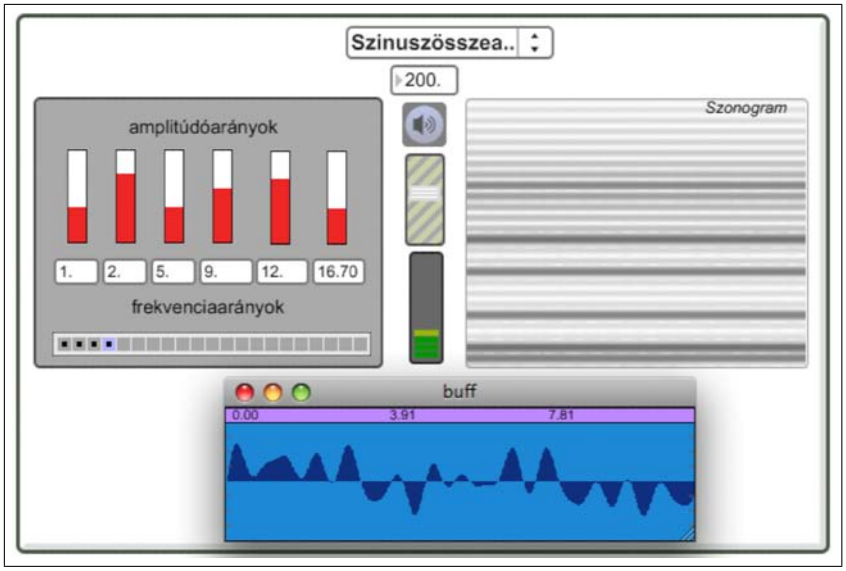
2.4. ábra. Szabadon szerkeszthető hullámforma. A bal oldali panel segítségével készíthetjük el saját hullámformánkat.

lefutása alatt több hullámformát alkalmaznak, egyiket a másikba úsztatva. Ezzel a módszerrel a szintézis még mindig nagyon hatékony marad, miközben a gépies hangszín gazdagabbá válik.

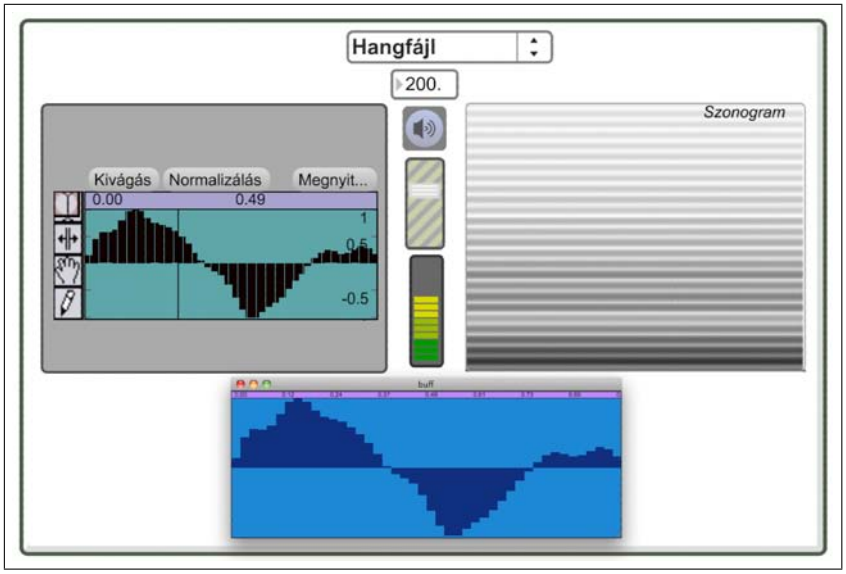
Paraméterek. A frekvencia és amplitúdó mellet a hullámformákat létrehozó módszerek paraméterei változtathatóak. A szinuszhullámok összeadásához szükséges paraméterek mennyisége általában túl nagy ahhoz, hogy azok könnyen átláthatóak legyenek. A mintavételezett hullámformáknál ez a probléma természetesen nem merül fel.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



2.5. ábra. Hullámforma létrehozása szinuszok összeadásával. Figyeljük meg, hogy a nem-egész arányok következtében a kapott hullámforma a periódus végén megszakad, emiatt a hangban lesz egy kattanás, torzulás. A szonogramban így sokkal több szinusz jelenik meg, mint amennyit az összeadáshoz felhasználtunk.



2.6. ábra. Mintavételezett hullámforma létrehozása.

Műveletigény. Az eljárás erőforrásigénye elhanyagolhatóan alacsony, leginkább ennek köszönheti népszerűségét.

2.4. Feladatok

1. A `HApp_02_01` `(OSX)` `(WIN)` minden lehetséges hullámforma-generálási módszerével hozz létre legalább egy-egy hangot!

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

2. Reprodukáld a 2.9, a 2.10 és a 2.11. hangpéldákat a `HApp_02_01` segítségével, vagy a `hd_02-01_csound.csd` megfelelő átírásával! Ellenőrizd az eredményt szonogram segítségével!

2.5. Függelék

2.1. CSound-forráskód. `hd_02_01_csound.csd`

```
<CsoundSynthesizer>
<CsInstruments>

sr      = 44100
kr      = 4410
ksmps   = 10
nchnls  = 1

instr 1
; fáziseltolás és táblázatkiolvavás
        kenv        linseg      0, .05, p4, p3-.1, p4*.8, .05, 0
        aphase      phasor      p5
        ilength     =           ftlen(p6)
        aphase      =           aphase*ilength
        asig        tablei      aphase, p6
        out         asig*kenv
endin


instr 2
; oszcillátor hullámforma megadásával
        kenv        linseg      0, .05, p4, p3-.1, p4*.8, .05, 0
        asig        oscili      kenv, p5, p6
        out         asig
endin

instr 3
; oszcillátor hullámforma megadásával, burkológörbe hullámforma megadá
    sával
        kenv        oscil1      0, p4, p3,08
        asig        oscili      kenv, p5, p6
        out         asig
endin

</CsInstruments>
<CsScore>

; 8192 mintából álló szinuszhullám
        f    01   0    8192     10   1

; négyszög  páratlan harmonikusokkal
        f    02   0    513 10   1 0 .333 0 .2 0 .143 0 .111 0 .0909 0 .077
             0 .0666 0 .0588

; fürész szinuszösszetevökböl
```

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02


```
        f    03  0    513 10  1 .5 .333 .25 .2 .166 .143 .125 .111 .1
            .0909 .0833 .077

; impulzussor
        f    04  0    4096    11  10  1 .9

; fürész
        f    05  0    512     7   1 512 -1

; négyszög
        f    06  0    1024    7   1 512 1 0 -1 512 -1

; egy impulzus
        f    07  0    128     7   1 1 1 0 0

; exponenciális burkoló
        f    08  0    1024    5   .001 10 1 824 .25 190 .001

        i 1     00  1 20000 440 01
        i .     02  .     .   . 02
        i .     04  .     .   . 03
        i .     06  .     .   . 04
        i .     08  .     .   . 05
        i .     10  .     .   . 06
        i .     12  .     .   . 07
        i .     14  .     .   . 08
        i .     15  .     0   . 08
        s

        i 2     00  1 20000 440 01
        i .     02  .     .   . 02
        i .     04  .     .   . 03
        i .     06  .     .   . 04
        i .     08  .     .   . 05
        i .     10  .     .   . 06
        i .     12  .     .   . 07
        i .     14  .     .   . 08
        i .     15  .     0   . 08
        s

        i 3     00  1 20000 440 01
        i .      +  .     .   . 02
        i .      +  .     .   . 03
        i .      +  .     .   . 04
        i .      +  .     .   . 05
        i .      +  .     .   . 06
        i .      +  .     .   . 07
        i .      +  .     .   . 08
        s

        i 3     00  0.5 20000  440 01
        i .      +    .     .  660 02
        i .      +    .     .  300 03
        i .      +    .     .  500 04
        i .      +    .     .  760 05
        i .      +    .     .  670 06
```

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02


```
99          i .       +     .      . 1000 07
100         i .       +     .      .  540 08
101         s
102
103 </CsScore>
104 </CsoundSynthesizer>
```

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

3. fejezet

Additív szintézis

Az additív szintézis a 2. fejezetben bemutatott hullámforma-szintézis általánosítása, mely során egymástól független (akár időben változó paraméterekkel bíró) szinuszhullámokat adunk össze.

3.1. Elmélet

A hullámforma-ismétlést továbbgondolva rájöhetünk, hogy a szinuszhullámok összeadása sokkal több lehetőséget rejt, mint egy szabályos hullámforma kikeverése. Amennyiben nem csak egy alaphang szabályos felhangjaival dolgozhatunk, olyan hangokat hozhatunk létre, melyek ugyan továbbra is összeolvadnak egyetlen hangszínné, mégis fluktuálnak. Egyszerű példaként gondoljunk egy eredetileg szabályos felhangokból felépülő hangszínre, ahol minden egyes komponens frekvenciáját pár centtel megváltoztatjuk. E megoldások, habár továbbra is összeolvadó hangszínt biztosítanak, *fényesebbé*, *élesebbé*, zeneibbé tehetik a végeredményt. Másrészt lehetőségünk nyílik olyan, ún. inharmonikus hangok megalkotására, amelyek ugyan nem zajszerűek, de nem is rendelkeznek önálló és egyértelmű hangmagassággal (akár azért, mert egyszerre több felhangsor hangjaból állnak, akár, mert elszórt komponenseket is tartalmaznak, vagy számos egyéb okból).

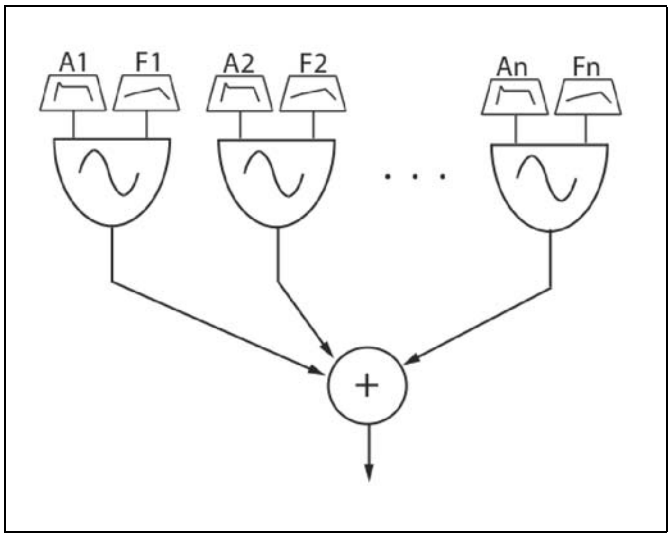
A szinuszhullámok szabad összeadásán alapuló módszer, az additív szintézis folyamatábrája a 3.1. ábrán látható. A hullámforma-szintézishez képest a két legfontosabb eltérés, hogy egyrészt nem csak egy adott frekvencia felhangjait tudjuk keverni, másrészt a frekvenciák és amplitúdók időben változhatnak. E változásokat *burkológörbékkel* írjuk le (ezeket általában töréspontok szakaszos összekötésével határozzuk meg, lásd a 3.2. ábrát). A töréspontokkal meghatározott burkológörbék és a véletlenszerűen fluktuáló periodikus változások (tremolo, vibrato) definiálásához felhasználható adatok mennyisége tetszőleges – összetett hangzások analíziséből nyert adatok reprodukálása esetén akár közelíthet a végtelenhez is. A hang részletessége, ami arányos az alkalmazott adathalmaz komplexitásával, egyedül a felhasználón múlik. E módszerrel bármilyen hangszín pontosan modellezhető, ám mivel az akusztikus hangok szerkezete rendkívül összetett, jelentős kihívást jelent természetesnek ható hangzások létrehozása.

Hogyan befolyásolják a hangszínt a választott frekvenciák és azok amplitúdói?

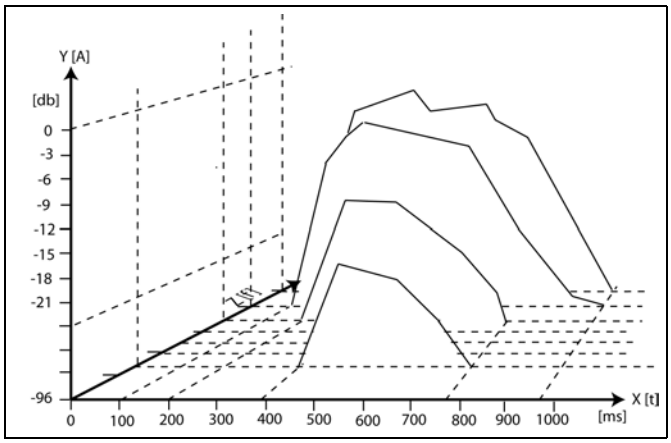
Az 1. fejezetben utaltunk rá, hogy bizonyos esetekben a tudatunk egy komplex hullámot egyetlen, meghatározott alaphanggal rendelkező hangszínnek fog fel: a spektrumot felépítő szinuszos komponensek ilyenkor „összeolvadnak”. Habár nem tudjuk az összeolvadás

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



3.1. ábra. Additív szintetizátor sematikus modellje, ahol minden oszcillátor (időben változó) frekvenciája és amplitúdója egyedileg szabályozható.



3.2. ábra. Egy trombitahang spektrális összetevőinek burkológörbéi.

pontos kritériumait, a következő megfontolások segítségünkre lehetnek:

Spektrum-összetevők aránya: a harmonikus összetevők olvadnak össze a legkönnyebben.

Kiegyensúlyozott frekvencia-eloszlás: ha a harmonikus összetevők túl messze helyezkednek el egymástól (akár egyenként, akár kis csoportokat alkotva), a hangzás szétszakadhat.

Kiegyensúlyozott hangerő-eloszlás: a kiugróan hangos összetevők egyenként hallhatóvá válnak. A „kiugró hangosság” megállapításához figyelembe kell vennünk, hogy a fül nem egyformán reagál különböző tartományokban. Nagyjából 1 és 3 kHz között kifejezetten jól hallunk, míg ettől lényegesen lejjebb, illetve feljebb sokkal nagyobb amplitúdók kellenek ahhoz, hogy egy szinuszhang azonos hangosságérzetet keltsen.

Összetevők kezdőpontjának szinkronicitása: ha az összetevők különböző időpontokban indulnak, a szinkronicitás megszűnik, az összetevőket egyenként is lehet hallani.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

Hasonló amplitudó-burkológörbék: ha az egyes összetevők dinamikai változása az időben hasonló, az segíti a fúziót. Ellenkező esetben az egyes összetevők hallhatóvá válnak.

Hasonló frekvencia-burkológörbék: ha az összetevők frekvenciája egymástól függetlenül változik, jól elkülönülnek egymástól (pl. vibrato segítségével jól elkülöníthetőek egymástól az összetevők).

A legtöbb hangolt hangszer spektruma egy adott alaphang felhangjait tartalmazza úgy, hogy a növekvő sorszámú felhangok amplitúdói csökkenő tendenciát mutassanak. Az, hogy melyik felhangokat hagyjuk ki (pl. csak páratlan felhangokkal dolgozunk, vagy mindegyikkel), alapvetően képes befolyásolni a hangszínt.

Habár eddig azt a benyomást keltettük, hogy forrásként mindenképp szinuszhangokat kell használnunk, ez koránt sincs így. Noha a „klasszikus" eljárás valóban szinuszokra épül, tetszőleges hullámforma szolgálhat az additív szintézis alapjául. Ez, bizonyos esetekben, egyszerűsíti a módszer használatát. Gyakori választás valamilyen ismert, egyszerű hullám (pl. fűrészfog-jel, négyszögjel stb.), de természetesen tetszőleges, akár egyedi tervezésű hullámforma is alkalmas. Ám amennyiben nem szinuszgenerátorokat használunk, számolnunk kell az alap-hullámformák által tartalmazott felhangokkal is.

Additív szintézis során gyakori megoldás az egyes oszcillátorok *csoportosítása*. Ilyenkor a teljes szintézis alapjául „additív mini-szintetizátorok" szolgálnak, amelyek maguk is több oszcillátort tartalmaznak. Ezeket az oszcillátor-csoportokat csupán néhány paraméterrel vezéreljük, az egyes források belső arányait előre rögzítjük. Ez történik a zongora magas hangjainál is: három – egymáshoz közeli frekvenciára hangolt – húr felelős minden egyes hangért. Az eljárás értelme, hogy az oszcillátor forráshangjai önmagukban is dús, organikus hangszínű jelek legyenek, ami kihat a teljes hangzásra.

3.2. Megvalósítás

Két (egy 8 és egy 48 csatornás) példán keresztül mutatjuk be az additív szintézist. Az egyszerűbb konfigurációt a `HApp_03_01` `(OSX)` `(WIN)`, a komplexebbet a `HApp_03_02` `(OSX)` `(WIN)` program mutatja be.

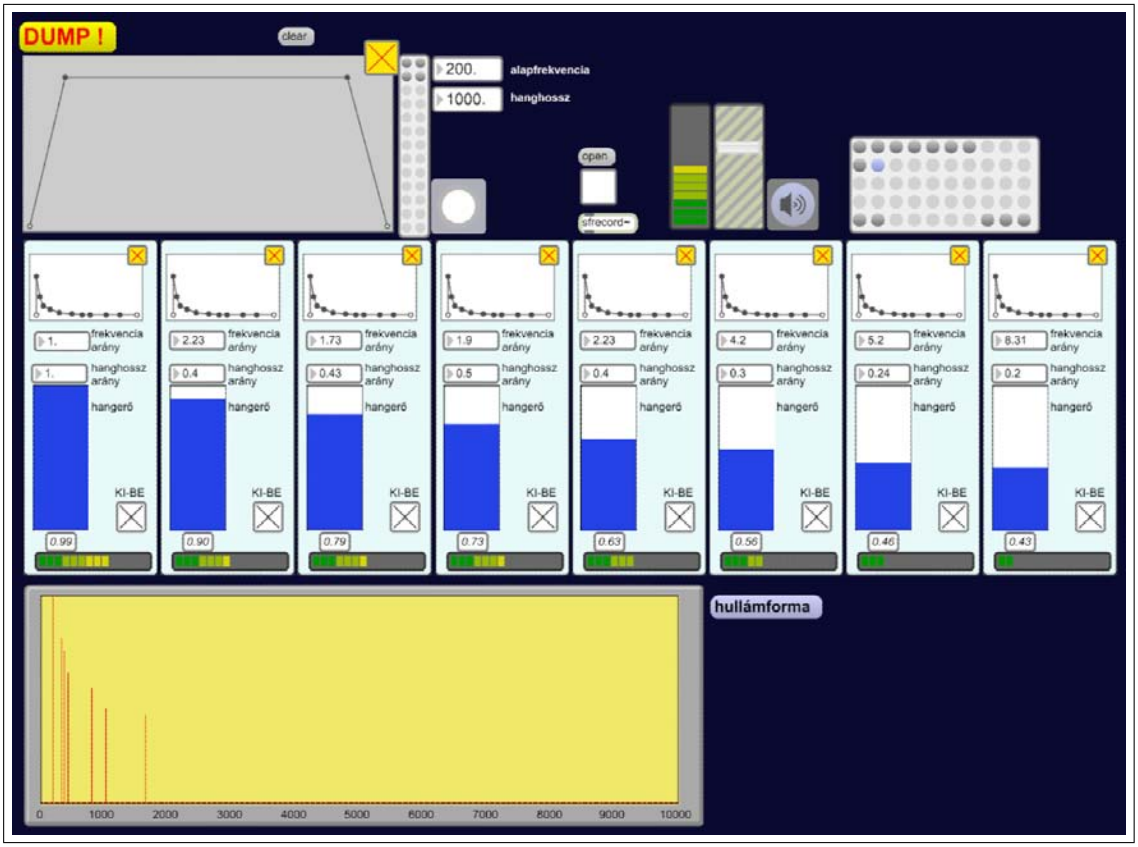
A 3.3. ábra középső sora tartalmazza a nyolc oszcillátor paramétereit egy-egy oszlopban. Lehetőség van az amplitúdó-burkolók egyenkénti megadására, illetve (a bal felső sarokban található szerkesztő segítségével) egységes beállításukra is. A hangerő-, időtartam- és frekvencia-arányokat a megfelelő csúszkák és számdobozok segítségével állíthatjuk be. Az időtartam és a frekvencia, amihez ezen arányok viszonyulnak, a felső sor közepén adhatók meg, az alattuk található gomb indítja a hangot. A jobb oldalon számos beállítás rögzítettünk előre; a képernyő alján a spektrumot láthatjuk.

A programban előre rögzítettünk pár beállítást.

A legfelső sorban (a balról második beállítást leszámítva) minden elmentett hangnak ugyanaz a spektruma: a beállított alaphang első 8 részhangja – csökkenő amplitúdókkal – , vagyis egy harmonikus hang. Egyedül az egyes komponensek időbeli lefutása változik. A balról 1. és 3. hang teljesen összeolvad, a 4. és 5. hang az eltérő burkológörbék miatt szétesik (ezen belül is az 5. beállítás egy dallamot hoz létre). A 6. beállítás annyiban tér el az 5.-től, hogy jóval gyorsabban megy végbe, így – habár hasonló dallamot alkotnak az egyes

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



3.3. ábra. Nyolc csatornás additív szintetizátor. A csatornák időtartam-, frekvencia- és amplitúdó-arányai, valamint amplitúdó-burkológörbéi egyedileg szabályozhatóak.

komponensek – az események túl gyorsak ahhoz, hogy valóban szétválasszuk őket (habár halljuk, hogy a hang nem egységes). A 7. beállítás egy hangszíndallamot hoz létre, ahol ugyan végig összeolvadó hangot hallunk, de ennek hangszíne folyamatosan (és folytonosan) változik. A 2. beállítás annyiban tér el az elsőtől, hogy a 2. és 3. oszcillátor inharmonikusan van hangolva, amitől a hangzás nem is olvad össze.

A második sorban eltárolt két hang spektruma szintén azonos, mégpedig egy erősen inharmonikus spektrum. Habár a balról legelső beállítás nem is olvad össze, a második (a komponensek azonos burkológörbéje és időtartama miatt) már összeolvad (lásd a 3.1. hangpéldát).

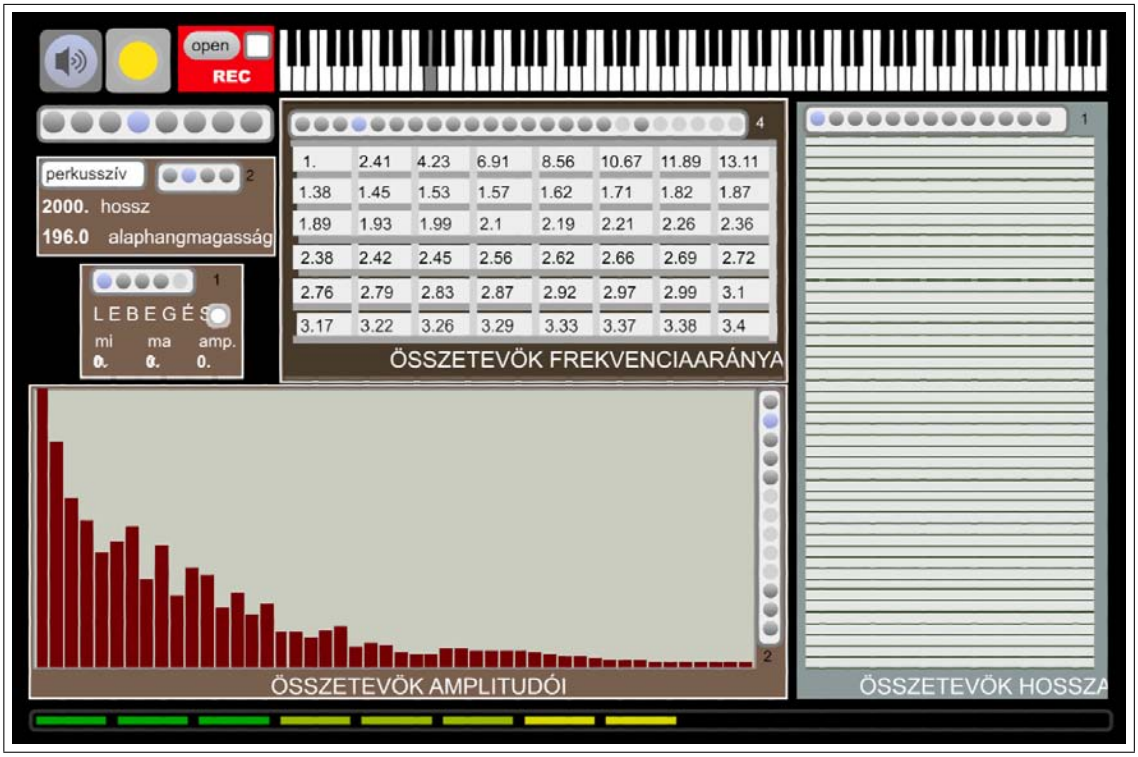
A másik példa ennél jóval bonyolultabb. Ez a rendszer 48 oszcillátorcsoportot tartalmaz, melyek mindegyike 3 szinuszgenerátorból épül fel. A 3.4. ábra a központi panelt mutatja, melyen a 48 csoport amplitúdóit, valamint az alapfrekvenciát, teljes időtartamát, az összetevők amplitúdó-burkológörbéjét és az oszcillátorcsoportokon belüli három generátor egymáshoz képesti lebegését lehet hangolni.

Az egyes összetevők frekvenciáit a 3.5. ábrán látható külön ablakban lehet megadni az alaphangmagasság frekvencia-értékének arányszámaival.

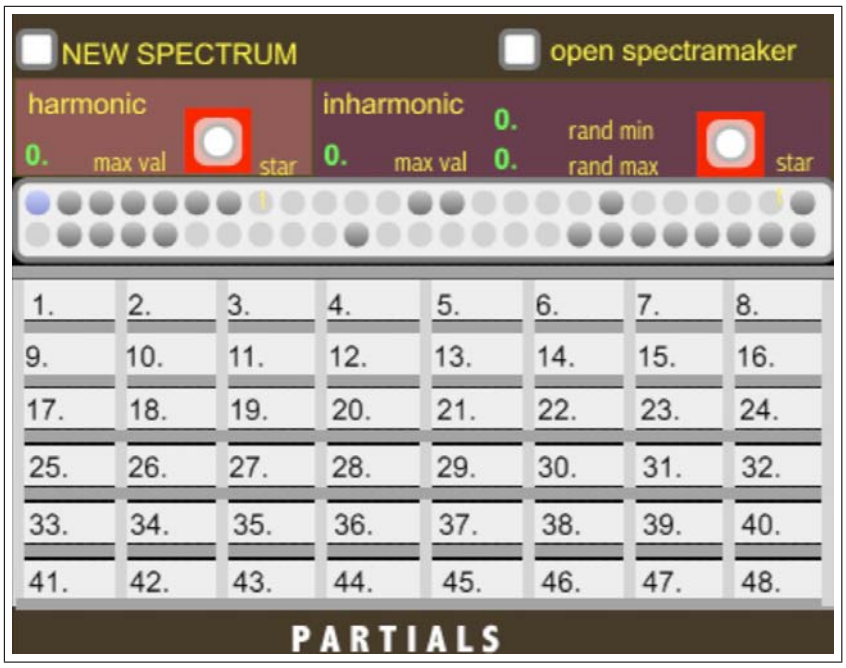
Az összetevők hosszúsága a főablakban megadott hanghossz szorzataival definiálható. Külön grafikus kezelőfelület alkalmazásával is bevihetjük az adatokat, ilyenkor vizuális visszacsatolást is kaphatunk az összetevők időtartamairól (lásd 3.6. ábra). Ez különösen jól működik ütött hangok esetén, amikor a hosszabb összetevők rezonanciaérzetet okoznak.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



3.4. ábra. Negyvennyolc csatornás additív szintetizátor.



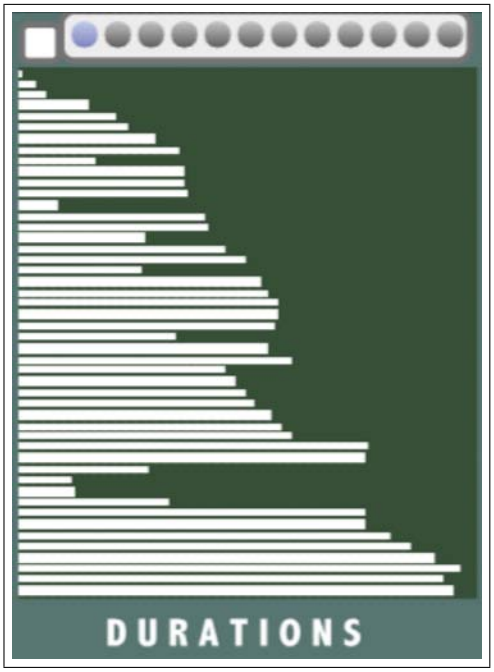
3.5. ábra. A frekvencia-arányok megadását szolgáló modul.

A grafikus ábra segítségével jól megfigyelhető, hogy egyes anyagszerű hangzásoknak (pl. cserép, üveg, stb.) hol vannak rezonanciahelyei. A burkológörbe típusától függően a „*hossz*” paraméter befolyásolhatja az összetevők összeolvadását, illetve szétesését is.

A programban előre rögzítettünk pár beállítást. Ezek egy része csak egy adott paraméter-

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



3.6. ábra. A komponensek időtartamait meghatározó szerkesztőfelület.

csoportot (pl. amplitúdó-arányok) vezérel, ám a főpanel (lásd a 3.4. ábrát) bal felső sarkában találhatóak teljes hangokat írnak le.

Az első három (illetve a balról legutolsó, lásd a 3.2. hangpéldát) beállítás egy-egy harmonikus hangot tartalmaz, melyek spektrumának frekvencia-arányai megegyeznek, ám míg az első és az utolsó hang összes komponensének amplitúdója egységnyi, a másik kettőben az amplitúdók csökkenő tendenciát mutatnak. A harmadik hangban az oszcillátor-csoportok úgy vannak hangolva, hogy az egyes komponensek lebegjenek (lásd a 3.3. hangpéldát). A legutolsó hangban az egyes komponensek időtartama eltérő, emiatt a spektrum időben változik (de összeolvad). A balról 4. és 5. hang (lásd a 3.4. hangpéldát) összeolvadó, de inharmonikus, perkusszív hangokat mutat be. A 6. és 7. hangok – elsősorban az eltérő időtartamok miatt – nem olvadnak össze (lásd a 3.5. hangpéldát).

A 3.1. CSound program egy speciális problémakört, a hangszín összeolvadásának a szinkronicitástól való függését vizsgálja Jean-Claude Risset szintetikus harangján keresztül. A harang hangja 9 szinuszhullámból tevődik össze, melyek időtartama mind különböző. Risset az egyes komponensek burkológörbéinek változtatásával kísérletezett; az eltérő időtartamok miatt így az azonos alakú görbék maximális amplitúdói eltérő időpontokba kerülnek. A csúcshelyek változtatásával a teljesen összeolvadó haranghang és a szinuszos komponenseire egyenként széteső hangszín közötti tetszőleges átmenet érhető el. A folyamatot a 3.6. hangpélda illusztrálja.

3.3. Összegzés

Hangszín. Nem létezik olyan hang, amit ne lehetne a módszerrel megalkotni. Kellően sok oszcillátor esetén a frekvenciák és az amplitúdók megfelelő időbeli paraméterezésével a legfinomabb hangszínfolyamatok is létrehozhatók, az eljárás által produkált hangszí-

Open Access 2026-11-02

nek a valódi hangszerekével is felveszik a versenyt. Ez azonban csak akkor igaz, ha kellő mennyiségű oszcillátorral és paraméterrel dolgozunk.

Paraméterek. A jó minőségű szintézis sok oszcillátort igényel, ám az ehhez felhasznált adatok mennyisége nem alkalmas arra, hogy kellően átlássuk (és vezérelhessük) őket: egy adott számú paraméter felett szinte már egyszerűbb lenne a létrehozandó minta közvetlen megrajzolása is. Ennek kiküszöbölésére különféle adatredukciós eljárások (pl. oszcillátorok csoportosítása) léteznek. Ha azonban túlságosan redukáljuk a paramétereinket, a hangszín könnyen gépiessé, élettelenné válhat.

Műveletigény. Túl sok oszcillátor egyidejű futtatása valós időben nem (vagy csak kivételesen erős számítógépeken) lehetséges, ám az oszcillátorok számának csökkentése könnyen a hangszín rovására mehet.

3.4. Feladatok

1. Jellemezd a `HApp_03_01` (OSX) (WIN) legalsó sorában levő beállításokat a harmonicitás, illetve összeolvadás szempontjából!

2. Kísérletezz a `HApp_03_02` (OSX) (WIN) beállításaival! Az egyes paramétercsoportokhoz rendelkezésre álló beállítások segítségével állíts össze legalább öt hangot!

3. Készíts 5 harmonikus és 5 inharmonikus hangot! Az így elkészített hangokból válassz ki kettőt-kettőt, és készítsd el őket legalább 3 különböző változatban, ahol a frekvencia- és hangerő-arányok azonosak, de a burkológörbéik eltérőek.

3.5. Függelék

3.1. CSound-forráskód. `hd_03_01_csound.csd`

```
<CsoundSynthesizer>
<CsInstruments>

sr      = 44100
kr      = 4410
ksmps   = 10
nchnls  = 1

instr 1
         idur1       =           p3
         idur2       =           p3*.9
         idur3       =           p3*.67
         idur4       =           p3*.55
         idur5       =           p3*.325
         idur6       =           p3*0.1
         idur7       =           p3*.25
         idur8       =           p3*.2
         idur9       =           p3*.15
         idur10      =           p3*.1
```

Open Access 2026-11-02


```
        idur11      =           p3*.075

        iamp1       =           p4
        iamp2       =           p4*0.67
        iamp3       =           p4
        iamp4       =           p4*1.8
        iamp5       =           p4*2.67
        iamp6       =           p4*1.67
        iamp7       =           p4*1.46
        iamp8       =           p4*1.33
        iamp9       =           p4*1.33
        iamp10      =           p4
        iamp11      =           p4*1.33

        ifqc1       =           p5*0.56
        ifqc2       =           p5*0.56+1
        ifqc3       =           p5*0.92
        ifqc4       =           p5*0.92+1.7
        ifqc5       =           p5*1.19
        ifqc6       =           p5*1.7
        ifqc7       =           p5*2
        ifqc8       =           p5*2.74
        ifqc9       =           p5*3
        ifqc10      =           p5*3.76
        ifqc11      =           p5*4.07

        ke1         oscil1      0, iamp1, idur1, p6
        ke2         oscil1      0, iamp2, idur2, p6
        ke3         oscil1      0, iamp3, idur3, p6
        ke4         oscil1      0, iamp4, idur4, p6
        ke5         oscil1      0, iamp5, idur5, p6
        ke6         oscil1      0, iamp6, idur6, p6
        ke7         oscil1      0, iamp7, idur7, p6
        ke8         oscil1      0, iamp8, idur8, p6
        ke9         oscil1      0, iamp9, idur9, p6
        ke10        oscil1      0, iamp10, idur10, p6
        ke11        oscil1      0, iamp11, idur11, p6

        a1          oscili      ke1, ifqc1, 1
        a2          oscili      ke2, ifqc2, 1
        a3          oscili      ke3, ifqc3, 1
        a4          oscili      ke4, ifqc4, 1
        a5          oscili      ke5, ifqc5, 1
        a6          oscili      ke6, ifqc6, 1
        a7          oscili      ke7, ifqc7, 1
        a8          oscili      ke8, ifqc8, 1
        a9          oscili      ke9, ifqc9, 1
        a10         oscili      ke10, ifqc10, 1
        a11         oscili      ke11, ifqc11, 1

        out         a1+a2+a3+a4+a5+a6+a7+a8+a9+a10+a11

endin

</CsInstruments>
<CsScore>
```

© *Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola*

Open Access 2026-11-02


```
76
77          f1   0 4096  10  1
78          f2   0 4096   5  1     4 10000 4092  1
79          f3   0 4096   5  1  2048 10000 2048  1
80
81          ;      idur  iamp  ifqc
82          i1   0    20 3000   400 2
83          i1   5    10 5000   400 3
84          i1  10    20 3000   150 2
85          i1  10    20 4000   150 3
86          e
87
88  </CsScore>
89  </CsoundSynthesizer>
```

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

4. fejezet

Szubtraktív szintézis

Az eddig bemutatott szintézistechnikák arra épültek, hogy a hangot annak komponenseiből építsék fel. Eljárhatunk fordítva is: ha a cél-hangszínnél sokkal sűrűbb spektrumból indulunk ki, a forrásul használt hang összetevőit megszűrve eljuthatunk a kívánt végeredményhez. E módszer neve szubtraktív szintézis. Mivel az eljárás analóg megvalósítása viszonylag egyszerű és gazdaságos, a szintetizátorok hőskorában ez volt az egyik legelterjedtebb szintézistechnika.

4.1. Elmélet

A megvalósítás két fő egységre oszlik: a szűrni kívánt forrásjelre, illetve a szűrőre.

4.1.1. Forrásjelek

Forrásjelként összetevőkben gazdag jeleket tanácsos használni, hiszen kevés összetevő kevés választási lehetőséget szolgáltat a kivonáshoz, illetve a kihangsúlyozáshoz. A leggyakrabban használt szintetikus forrásjelek:

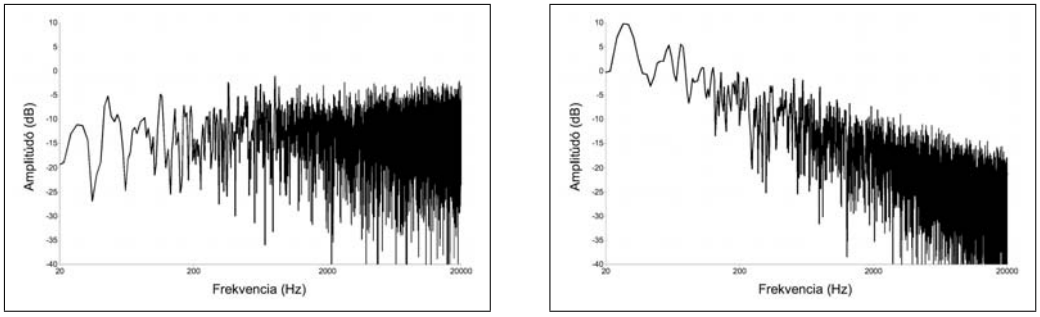
Széles spektrumú zajok. A teljes hallható tartományban tartalmaznak összetevőket. Kiváló források, hiszen minden szűrővel eredményesen lehet változtatni a hangzásukat, zajos és harmonikus hangok létrehozására egyaránt alkalmasak. Aszerint szokás csoportosítani őket, hogy a spektrumuknak megfelelő látható fény milyen színű. A legnépszerűbbek a *fehérzaj* (minden összetevő energiája azonos, lásd a 4.1. hangpéldát) és a *rózsazaj* (oktávonként 3 dB esés a spektrumban, lásd a 4.2. hangpéldát). Spektrumukat a 4.1. ábra mutatja.

Harmonikus jelek. Olyan periodikus jelek, amiknek gazdag a spektruma. Ezek kizárólag harmonikus összetevőket tartalmaznak, ezért konkrét hangmagassággal rendelkező hangzások kialakítására alkalmasak. A legnépszerűbbek a négyszög-, a háromszög- és a fűrészfogjel, valamint az impulzussorozat (lásd a 2. fejezetet).

Természetesen bármilyen előre rögzített vagy élőben beérkező hang is jó alapanyag lehet szűréshez, amennyiben összetevőkben gazdag spektrummal rendelkezik.

© *Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola*

Open Access 2026-11-02



4.1. ábra. Fehérzaj (bal) és rózsazaj (jobb) jellemző spektrumai. Logaritmikus frekvenciaábrázolás mellett (ekkor az azonos vízszintes távolság azonos hangköznek felel meg) jól látható, hogy míg a frekvenciák amplitúdója fehérzaj esetén gyakorlatilag állandó, a rózsazajban az amplitúdók csökkenő tendenciát követnek (oktávonként 3 dB esést).

4.1.2. Szűrőtípusok

A szűrőket a (frekvenciafüggő) *amplitúdó-* és *fázisválaszaik* határozzák meg. Előbbi azt mondja meg, hogy a beérkező jel spektruma hogyan változik meg a szűrő hatására, míg utóbbi azt adja meg, hogy a spektrum egyes frekvencia-komponensei mennyivel „tolódnak el" időben. A szűrőket elsősorban amplitúdó-válaszuk (más szóval *amplitúdó-karakterisztikájuk*) szerint csoportosítjuk. A legfontosabb szűrőtípusok a következők (lásd 4.2. ábrát):

Aluláteresztő: csak adott frekvencia alatt enged át összetevőket.

Felüláteresztő: csak adott frekvencia felett enged át összetevőket.

Sáváteresztő: csak egy adott középponti frekvencia körüli megadott sávszélességben enged át összetevőket.

Sávzáró: egy adott középponti frekvencia körüli megadott sávszélesség kivételével enged csak át összetevőket.

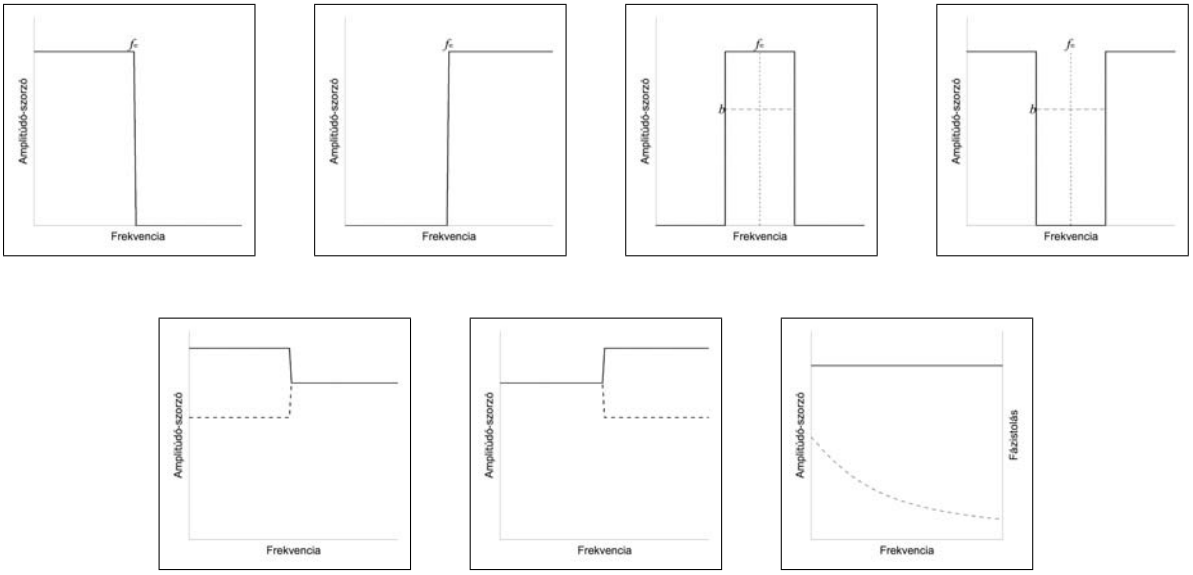
Shelf-szűrők: minden frekvenciát átenged, de egy adott frekvencia felett (high-shelf) vagy alatt (low-shelf) felerősíti vagy tompítja a spektrumot.

Mindent áteresztő: minden frekvenciát átereszt; hatása egyedül a fázis-karakterisztikában érhető tetten: mindent áteresztő szűrőket akkor alkalmazunk, amikor egyedül a beérkező jel fázisát szeretnénk megváltoztatni. A leggyakrabban használt mindent áteresztő szűrők egy adott frekvencia alatt változatlanul hagyják a fázisokat, míg az adott frekvencia felett 180 fokkal eltolják őket (felhasználásukat lásd a 12. fejezetben).

A szűrőket (a közép-, illetve a levágási frekvencia mellett) jól jellemzi a *jósági tényezőjük*: ez a szűrőből távozó és a szűrőbe beérkező jelek teljesítményeinek hányadosa (sávszűrők esetén jól közelíthető a középfrekvencia és a sávszélesség hányadosával). Általában minél nagyobb egy szűrő jósági tényezője, annál inkább korlátozott az a frekvencia-tartomány, amit a szűrő áteresztő magán. Ezért a nagy jósági tényezőjű szűrőket gyakran *rezonáns* szűrőknek hívjuk, hiszen az ilyen eszköz bármilyen bemenő jelből egyetlen – általában szűk – frekvencia-tartományt emel csak ki.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



4.2. ábra. A legfontosabb szűrőtípusok idealizált amplitúdó-karakterisztikái. Felső sor (balról jobbra): aluláteresztő, felüláteresztő, sáváteresztő, sávzáró; alsó sor (balról jobbra): low-shelf, high-shelf, mindent áteresztő. f_c (típustól függően) a levágási- vagy a középfrekvenciát, míg b a sávszélességet jelöli. A valóságban a karakterisztikák levágása sosem végtelenül éles; a levágási tartomány meredeksége *fontos jellemzője az adott szűrő minőségének. A mindent áteresztő szűrő esetén (mivel amplitúdókarakterisztikája teljesen vízszintes) feltüntettük a fázis-karakterisztikát is.*

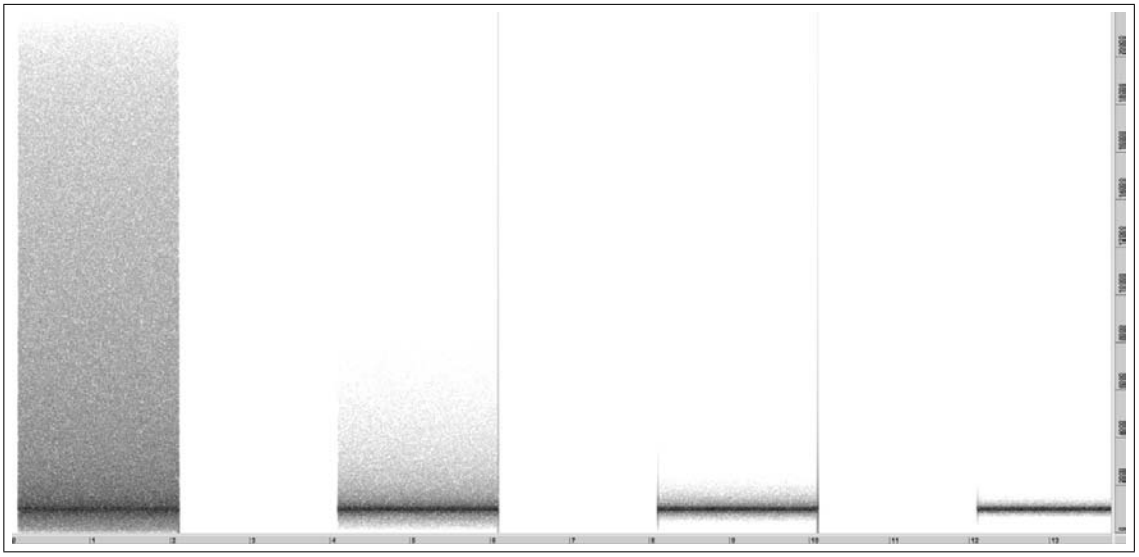
Sáváteresztő szűrőket sorosan (egymás után) és párhuzamosan is lehet alkalmazni (utóbbi konstrukciót *ekvalizátornak* hívjuk). Soros kötés esetén meredekebb lesz a szűrő, így jobban el lehet különíteni a kívánatos spektrumrészt (lásd a 4.3. ábrát). Párhuzamos kötéskor az additív szintézis egy speciális esetét lehet modellezni: a spektrum hasonló lesz ahhoz, mint amikor az egyes szűrők középfrekvenciáit használva additív szintézist valósítunk meg, ám a szinuszok helyett (melyek a spektrumban egyetlen függőleges vonalként jelennek meg) itt kiterjedt sávok lesznek, ezáltal a hangzás vibrálóvá, bonyolultabbá válik. A spektrum összetevőinek aránya az egyes sávszűrők csillapítási mértékének változtatásával lehetséges. A technika egyik előnye, hogy viszonylag kevés paraméter változtatásával vizsgálhatóak az összetevőkben rendkívül gazdag zajos hang (széles sávszélesség esetén) és a szinuszos összetételű (keskeny sávszélesség) közötti területek. Az egyes sávok amplitúdójának és sávszélességének független kezelésével egyszerűen alakíthatóak ki rezonanciahelyek. A módszer használatának problémája, hogy a szűrt sávok hangereje általánosságban nem becsülhető meg előre (kivéve, ha pontosan ismerjük a beérkező szűrt jel spektrális összetételét), így összetett hangszínterek esetén a jelen idejű normalizálás nem mindig lehetséges. Ez nehézséget jelenthet élő játék esetén.

4.2. Megvalósítás

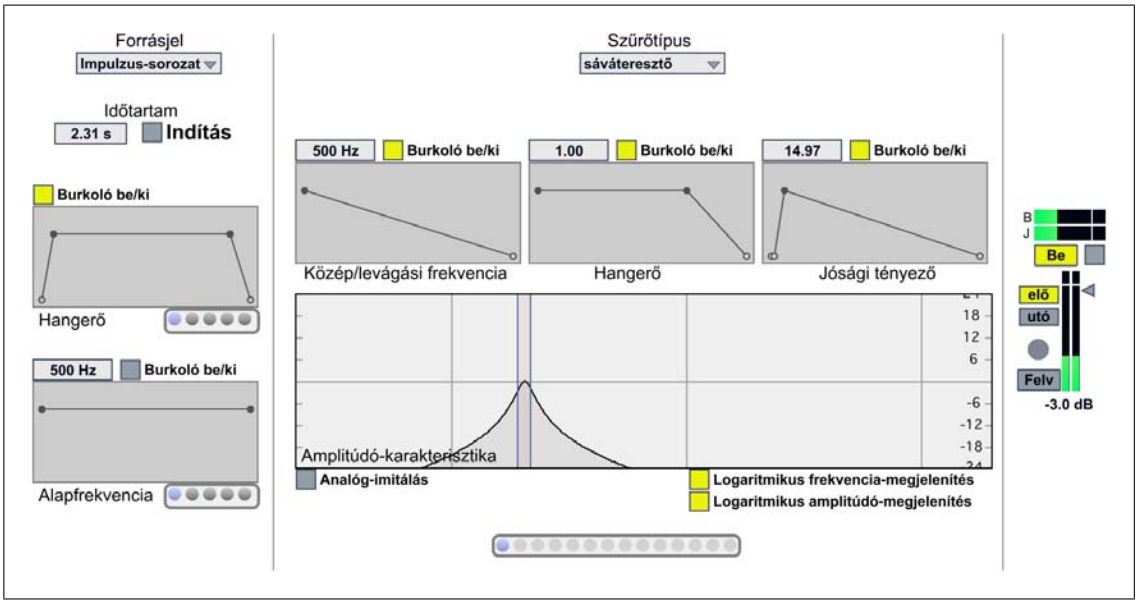
A `HApp_03_01` `(OSX)` `(WIN)` példaprogram mutatja be az alapvető forrásokat és szűrőtípusokat (lásd a 4.4. ábrát).

A program lehetőséget ad arra, hogy az egyes paramétereket burkológörbékkel szabá-

Open Access 2026-11-02



4.3. ábra. Fehérzaj szűrése több lépcsőben. A felhasznált alapszűrő egy sáváteresztő karakterisztikájú, 1 000 Hz középfrekvenciájú, 100 Hz sávszélességű szűrő. Az egyes lépésekben rendre 1, 2, 3 és 4 ilyen szűrőt kapcsoltunk össze soros elrendezésben. A szonogram függőleges tengelyén a frekvencia (Hz), vízszintes tengelyén az idő (másodperc) szerepel. Az eredményt a 4.3. hangpélda tartalmazza.



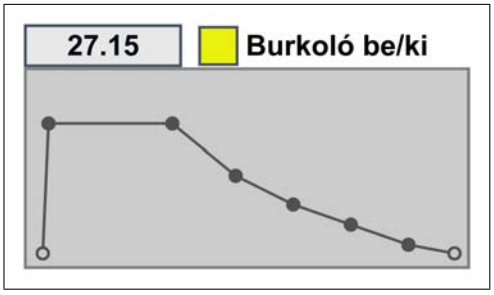
4.4. ábra. A `HApp_03_01` példaprogram. A jobb oldali szűrővel különböző forrásokat szűrhetünk.

lyozzuk. Ezeket a paramétereket úgy jelenítettük meg, hogy a konstans értéket meghatározó számdoboz mellé egy „*Burkoló*” feliratú kapcsolót, alá pedig egy burkológörbe-szerkesztőt helyeztünk (lásd a 4.5. ábrát).

A burkológörbék egységes időtartamát a „*Lejátszás*” számdoboz adja meg; a lejátszás az „*Indítás*” feliratra kattintva indul.

A bal oldali kezelőfelület a forrást vezérli. A lehetséges opciók: fehér- és rózsazaj, im-

Open Access 2026-11-02



4.5. ábra. Burkológörbével modulálható paraméter. A jobb felső „Burkoló” kapcsoló kikapcsolt állapotában a paramétert a bal felső számdoboz határozza meg. Bekapcsolt állapotban a bal felső számdobozban megadott konstans szorzódik a burkológörbével. Utóbbi értékei mindig 0 és 1 közé esnek.

pulzussorozat, háromszög-, fűrészfog- és négyszögjel. A „*Hangerő*” burkológörbe bekapcsolásával a hangforrás (egyébként konstans) hangerejét tudjuk modulálni. A nem-zajszerű forrásokhoz egy második paraméter is tartozik: az alapfrekvencia, melyet az alsó burkológörbével állíthatunk be.

A jobb oldali felület a szűrőt szabályozza. A szűrőtípus kiválasztása után megadhatjuk a közép-, illetve levágási frekvenciát (szűrőtípustól függően), valamint a hangerőt és a jósági tényezőt. A szűrő hangerejét (erősítését) szabályozó paraméternek csak abban az esetben lesz hatása, ha a felette található „*Hangerő engedélyezés*” kapcsoló aktív. Az „*Analóg imitálás*” opció bizonyos szűrőtípusok esetén úgy módosítja az amplitúdó-karakterisztikát, hogy az jobban hasonlítson az analóg szűrők karakterisztikáira.

A forrás amplitúdója és frekvenciája, valamint a szűrők főbb paraméterei külön-külön menthetőek és visszaolvashatóak.

A `HApp_03_02` `(OSX)` `(WIN)` program egy nyolcsávos ekvalizátor segítségével szűr fehérzajt (lásd a 4.6. ábrát). A felső sorban egy zongorabillentyűzet segítségével adhatjuk meg az „alaphangot”, ami az egyes sávok középfrekvenciáit a bal oldalon található nyolc szorzón keresztül határozza meg. A sávok amplitúdói és jósági tényezői (ezen keresztül a sávszélességei) külön-külön szabályozhatóak, de lehetséges ezeket előre rögzített beállítások alapján is betölteni. A hangerőre három különböző burkológörbe illeszthető:

Impulzusok: a hang teljes időtartama alatt véletlen számú és elhelyezkedésű impulzus formájában jön létre a hang.

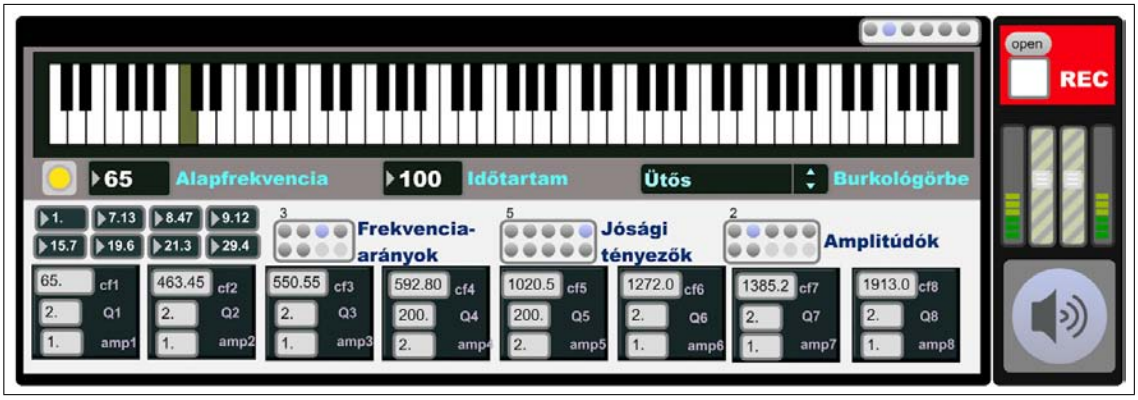
Kitartott hang: a hang egyenletes hangerővel szól annak teljes időtartama alatt.

Ütős: gyorsan lecsengő hangerő-burkológörbe, ami leginkább az ütőhangszerek hangjaira jellemző.

A programhoz számos beállítást készítettünk.

A balról 1., 4. és 5. hangban a középfrekvenciák harmonikus rendben követik egymást, a jósági tényezők pedig (egy-egy kivételtől eltekintve) nagyon magasak (vagyis a sávokhoz tartozó sávszélességek nagyon szűkek), így a végeredmény az additív szintézisre emlékeztet. A hangok között két alapvető különbség van. Egyrészt az 1. hang kitartott burkológörbét, míg a másik kettő perkusszív jellegűt használ. Emellett a 4. beállítás (a másik két említettel ellentétben) az alsó két, illetve a legmagasabb csatornán lényegesen kisebb jósági tényezőket (vagyis nagyobb sávszélességeket) ír elő, így ez a hang jóval zajosabb lesz, mint a másik

Open Access 2026-11-02



4.6. ábra. Fehérzajt szűrő nyolcsávos ekvalizátor.

kettő (lásd a 4.4. és a 4.5. hangpéldákat). A 3. beállítás szintén egy harmonikus hangot takar, ám itt a frekvencia-arányok mind páratlanok, amitől a hangszín zajjal vegyített klarinétszerű jelleget ölt. A 2. és a 6. beállítás inharmonikus hangokat eredményez, ám míg a 2. egy szélessávú, megütött zajt ad (mely nagyon hasonlít a megpendített üveghez, lásd a 4.6. hangpéldát), a 6. hang egy nagyon szűk sávra koncentráló, szél-jellegű morajlásra emlékeztet.

A 4.1. és a 4.2. CSound-programok különböző szűrési eljárásokat ismertetnek. A 4.1. példában egy rezonáns szűrővel szűrünk fehérzajt (az eredményt a 4.7. hangpélda tartalmazza), mind konstans, mind változó időbeli paraméterekkel. A 4.2. program ellenben egy pengetett hangforrást ereszt át két, párhuzamosan kapcsolt szűrőből (egy aluláteresztő és egy sáváteresztő) álló rendszeren (lásd a 4.8. hangpéldát).

4.3. Összegzés

Hangszín. Nagyon sokféle hang létrehozható a segítségével annak függvényében, milyen forráshangokat és milyen szűrőket használunk. Mivel működése modellezi az emberi hang és számos hangszer megszólalásának módját, alkalmas ezek létrehozására is (lásd a 7. fejezetet). Sőt, mivel a párhuzamos sávszűrés (ekvalizátor) szélsőséges módja (nagyon sok sáv nagyon keskeny sávszélességgel) az additív szintézis egyik formájának tekinthető, elvileg végtelen mennyiségű hangszín modellezhető segítségével. A paraméterek időbeni változtatásával sok összetevő mozgatható egyszerű módon, ami természetközeli hangzások keltésére is alkalmas. Az egyszerűbb szűrőkkel működő szintetizátorok és szoftverek természetesen korlátozzák a szintézistechnika lehetőségeit.

Paraméterek. A szűrés paramétereinek hatását egyszerű megérteni, érzékelni és vezérelni. A törésfrekvencia, a középfrekvencia, a sávszélesség és a sávok amplitúdóaránya intuitív módon kezelhető. A szűrés paramétereinek változtatása folyamatos érzetet biztosít, diszkrét pontjaikból egyszerű skálákat építeni.

Műveletigény. Habár más „alapműveletekhez” (hullámforma-kiolvasás, késleltetés stb.) képest a szűrés viszonylag sok erőforrást igényel, mivel a szintézistechnika során kevés szűrőt szokás alkalmazni, a módszer erőforrásigénye összességében alacsony.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

4.4. Feladatok

1. Hallgasd meg a HApp_03_01 (OSX) (WIN) program által felkínált összes forrást! Válassz ki egy szűrőtípust és állítsd be a paramétereit (ne használj burkológörbéket)! Hallgasd meg a szűrőn keresztül az összes forrást! Ismételd meg az eljárást különböző szűrőkkel és beállításokkal, legalább 5-ször!

2. A HApp_03_02 (OSX) (WIN) egyes csoportjait vezérlő lementett beállítások felhasználásával hozz létre legalább 5 hangot. Legyen köztük harmonikus, inharmonikus, zajszerű, konkrét hangmagassággal rendelkező, megütött és kitartott is!

3. Hozz létre...

   - ...megütött fára...
   - ...megütött cserépre...
   - ...megütött üvegre...
   - ...szélre...
   - ...morajlásra, földrengésre...

   ...emlékeztető hangokat! Próbálj meg minél többet úgy megvalósítani, hogy a paramétereik az időben változó lecsengésűek (burkológörbével vezéreltek) legyenek!

4.5. Függelék

4.1. CSound-forráskód. hd_04_01_csound.csd

```
<CsoundSynthesizer>
<CsInstruments>

sr      = 44100
kr      = 4410
ksmps   = 10
nchnls  = 1

instr 1
        anoise      rand        20000
                                                ; fehérzaj
        kcfreq      expseg      p5, 2*p3/10, p5, 6*p3/10, p6, 2*p3/10,
           p6 ; középfrekvencia burkoló
        kbwdth      linseg      p7, 2*p3/10, p7, 6*p3/10, p8, 2*p3/10,
           p8 ; sávszélesség burkoló
        kenv        oscil1      0, p4, p3, p9
                                        ; amplitúdó burkoló
        afilt       reson       anoise, kcfreq, kbwdth, 2
                            ; rezonáns szűrő
        out         afilt*kenv
endin

</CsInstruments>
<CsScore>
```

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02


```
20
21          f1 0 2048 5 0.001  10 1 2038 0.001
22          f2 0 2048 5 0.001   2 1 2046 0.001
23          f3 0 2048 5 0.001 200 1 1648 1     200 0.001
24
25          i 1 0    5   0.5    100    100   100     1   3
26          i 1 +    5   0.5    500    500    10     1   3
27          i 1 +   10   0.5    100   1000     5     5   3
28          i 1 +    3   0.5    100    200   100   100   1
29          i 1 +    3   0.5   1000   1200    10     5   2
30          e
31
32  </CsScore>
33  </CsoundSynthesizer>
```

4.2. CSound-forráskód. `hd_04_02_csound.csd`

```
1   <CsoundSynthesizer>
2   <CsInstruments>
3
4   sr     = 44100
5   kr     = 4410
6   ksmps  = 10
7   nchnls = 1
8
9   instr 1
10          idur        =           p3
11          iamp        =           ampdb(p4)
12          ifrq        =           cpspch(p5)
13          icut1       =           p6
14          icut2       =           p7
15          asource     pluck       iamp, ifrq, ifrq, 0, 1 ; pengetett
               hangforrás
16          kcut        expon       icut1, idur, icut2     ; törésfrekvencia
                burkoló
17          abpf        butterbp    asource, kcut, kcut*.2 ; sávszűrő
18          alpf        butterlp    asource, kcut          ; aluláteresző sz
               űrő
19          amix        =           abpf*p8+alpf*(1-p8)
20          out         amix
21  endin
22
23  </CsInstruments>
24  <CsScore>
25
26          i 1 0 .5 120 8.09   200  100 1
27          i . + .   <    <      <    . 1
28          i . . .   .    .      .    . 1
29          i . . .  90 8.12   200    . 0.8
30          i . . .   <    <      <    . <
31          i . . .   . 8.01  3000 3000 0.5
32          i . . .   . 8.03     <  200 .
33          i . . .   . 8.06     <    < .
34          i . . .  70 8.09 10000  100 .
35          e
36
```

Open Access 2026-11-02


```
37 </CsScore>
38 </CsoundSynthesizer>
```

© *Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola*

Open Access 2026-11-02

5. fejezet

Granuláris szintézis

Az 1. fejezetben beszéltünk arról, hogy a „mikroszkopikus” időskálán a frekvencia – illetve maga a hangszín – nem függetleníthető az időbeliségtől. Pontosabban, a hang periódusidejének megfelelő (vagyis az ezredmásodperc körüli) nagyságrendben egyáltalán nem beszélhetünk külön időbeli és frekvenciabeli folyamatokról, és ahhoz, hogy világosan elkülöníthessük a frekvenciát és az időt, legalább egy nagyságrendet (századmásodperces tartomány) kell ugranunk. A jelenség (matematikailag) teljesen hasonló a fizikából ismert hullám-részecske problémához, talán nem véletlen, hogy épp egy fizikus (Gábor Dénes) ötlete nyomán jött létre a granuláris szintézis.

5.1. Elmélet

Az eljárás lényege, hogy „hangrészecskék”, más néven „*szemcsék*” segítségével építsünk fel hangokat, mint ahogy a valódi részecskék építik fel az anyagot. A szemcsék a hangzás „atomjai”: a kapott hangszínt e szemcsék belső felépítése (mikroszerkezet) és a szemcsék összekeverésének módja (makroszerkezet) együttesen határozzák meg. Az 5.1. ábra három eltérő szerkezetű szemcsét mutat.

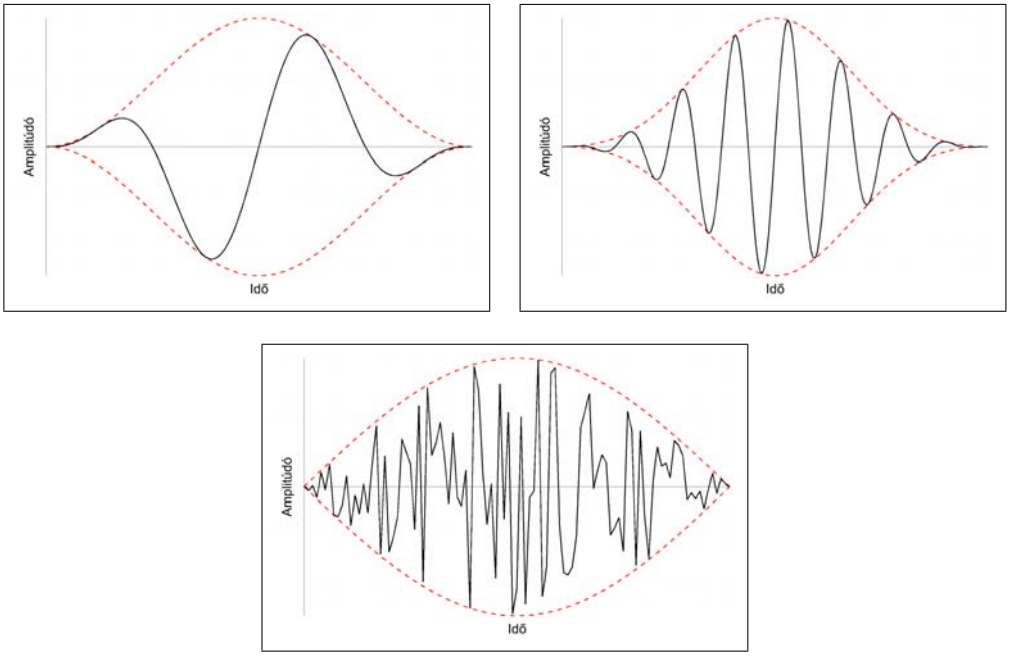
A szemcséket elkülöníthetjük felvett hangokból, de létrehozhatjuk szintézis útján is. A granuláris szintézis fő paraméterei:

Hullámforma. A spektrum elsősorban a szemcsét kitöltő hang hullámformájától függ, így ez határozza meg a hangszín alapvető jellegét. Gyakorlatilag bármilyen jelet fel lehet használni granuláláshoz, az egyszerű szinuszhullámtól az időben változó, komplex hangzásokig, így rendkívül széles a létrehozható hangzások palettája.

A hullámforma két jel szorzatából tevődik össze. Az egyik az a (végtelen időtartamúnak elképzelt) „idealizált jel”, amiből a szemcsét mintavételezzük (ha felvett hanggal dolgozunk, ez a jel maga a felvett hang). A másik az az amplitúdó-burkológörbe, amivel „kivágjuk” a szemcsét ebből az idealizált jelből. Utóbbi alakja határozza meg, hogy milyen összetevők adódnak hozzá a szemcsét alkotó hullámforma spektrumához. Ha a burkológörbe a szemcse elején és végén is nulla értéket vesz fel, valamint alakja „kerekded” (törésmentes), a spektrum nem fog sok, az eredeti hullámforma hangszínét akár a zajig is eltorzító összetevőt tartalmazni; ezzel szemben a négyszögletes, hirtelen változásokat tartalmazó burkológörbe zajos kopogásokkal tarkítja a granulált hangzást. Az eljárás a 10. fejezetben bemutatott *ablakolás* analógiája.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



5.1. ábra. Három különböző hangszemcse. A szaggatott vonalak az amplitúdóburkológörbéket mutatják, amiket a szemcsék „kivágásához" használtunk (e burkológörbéket a gyakorlatban valamilyen ablakfüggvénynek – lásd a 10. fejezetet – szokás választani). A két felső szemcsét szinuszos hangból generáltuk, a kettő közti hangszínkülönbség az eltérő belső szerkezetből, pontosabban, az oszcillációk eltérő számából adódik (lásd az 5.1. és az 5.2. hangpéldákat). A lenti szemcsét egy zajforrással hoztuk létre (lásd az 5.3. hangpéldát).

A granuláris hangzás fontos meghatározója, hogy az egyes szemcséket kitöltő hullámforma miként változik az időben. Amennyiben felvett hangból különítjük el a szemcséket, a következő módszereket alkalmazhatjuk:

- Időben nem változó hullámforma esetén „befagyasztott", konstans hangszín keletkezik (lásd az 5.4. hangpéldát).
- Az eredeti hangot (illetve egy ahhoz nagyon hasonlót) kapunk, ha a szemcséket a kiindulási hangban fokozatosan előrehaladva, folyamatosan mintavételezzük. Az így szintetizált hangot transzponálhatjuk és nyújthatjuk is (lásd az 5.5. hangpéldát).
- Ha a szemcsékhez egyenkét címezzük a hullámformákat, a hangzás mikroszkopikusan komponálhatóvá válik (lásd az 5.6. hangpéldát).
- A szemcséket egy adott időtartamon belül véletlenszerűen választva „összekeverhetjük" az eredeti hangot. Minél nagyobb az időtartam, amiből a szemcsék választhatóak, annál kevésbé lesz felismerhető az eredeti hang (lásd az 5.7. hangpéldát).

Szemcsehossz. Meghatározza, hogy milyen mértékben érvényesül a szemcsét kitöltő hullámforma hangszíne a granulált hangzásban. Ahogy a fejezet elején említettük, az időskála közvetlen hatással van a hangszínre, a következő percepciós kategóriahatárok szerint:

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

- 2 ms-nál rövidebb időtartam esetén nem lehet érzékelni a szemcsét alkotó hullámforma hangmagasságát, az egyes szemcsék zajos kattanásként szólnak. A hosszúságtól függően változik a szemcse kattanásának színezete.
- 5 ms körül elkezd kialakulni a hangmagasságérzet, ami ennél az értéknél még rendkívül bizonytalan, és függ a szemcsét kitöltő jel hangmagasságától.
- 25 ms felett egyre biztosabb hangmagasságérzet alakul ki.

A három határérték között elhelyezkedő területek átmenetet alkotnak a leírt állapotok között. Az egyes szemcsék hosszát folyamatosan és véletlenszerűen is lehet változtatni.

Szemcsefrekvencia. Megfelelő hosszúságú szemcsék esetén a hullámforma hangmagassága meghatározza a granulált hangzás hangmagasság-regiszterét.

A granuláris szintézis során bevett módszer a szemcsék transzpozíciója, melynek a gyakorlatban három fő módszere ismert:

- Az eredeti hang transzponálása a szemcsék egyformán történő transzponálásával érhető el (lásd az 5.5. hangpéldát).
- Mikroszkopikusan komponálhatunk hangmagasság-viszonyokat, ha a szemcséket egyenként megcímezve, külön értékekre transzponáljuk. Így különböző skálákon alapuló szemcsedallamok hozhatók létre.
- Ha a szemcséket véletlenszerűen transzponáljuk (adott minimális és maximális értékek között), a szélső értékek beállításától függően az apró remegésekkel töredezettől az intenzíven mozgó textúrákig lehet eljutni (lásd az 5.8. hangpéldát).

Szemcsesűrűség. A szemcsék követési gyakoriságát adja meg. E gyakoriság (frekvencia) hatása a hangra:

< 8 Hz: különálló megszólalások, nem alakul ki folyamatos hangzás (lásd az 5.9. hangpéldát).

8–15 Hz: gyorsan ismétlődő, pergő hangsorozat (lásd az 5.10. hangpéldát).

15–80 Hz: a szemcsét alkotó hanghullám tulajdonságaitól függően a terület fontos határvonal, ugyanis ez a paraméter határozza meg, hogy két egymást követő hang elkülönül-e egymástól, vagy egy hangeseménnyé olvad össze.

80 Hz <: a granulálás frekvenciája nagymértékben befolyásolja a keletkező hangzás minőségét, hasonlóan az amplitúdó-moduláció által okozott spektrumváltozáshoz. Különösen akkor nyilvánvaló a hatás, ha a szemcsék hossza nagyon rövid, megközelíti vagy eléri a zajos kattanások területét. Hosszabb szemcsék esetén komplex, dús hangszövetek keletkeznek, melyek tulajdonságai a szemcsék paramétereitől függnek (lásd az 5.11. hangpéldát).

A sűrűséget befolyásolja, hogy az adott szoftver hány szólamot képes egyszerre lejátszani.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

Lokalizáció. A szemcsék térbeli eloszlását befolyásolja. Sztereo formátumban általában a jobb és bal oldal közötti panoráma paramétert jelenti, de lehetséges különböző mélységérzetet is generálni az egyes szemcsékhez. A megfelelő megoszlásérzet érdekében szemcsénként más értékekkel kell ellátni a panoráma paramétert.

A granuláris szintézis több ponton is hasonlít az additív szintézishez, egyrészt általánossága (elvileg bármilyen hang szintetizálható a segítségével), másrészt a vezérlés bonyolultsága miatt, hiszen e módszer is rengeteg paramétert igényel. Az elvileg szemcséről szemcsére változó adatok mennyisége miatt egy másodpercnyi hang létrehozása is komoly problémát jelentene, ha egyenként kellene megadni e változókat. Szerencsére a granuláris szintézis során a kívánt hangszíneket kis mennyiségű, általános tulajdonságokat leíró paraméterek segítségével is létrehozhatjuk. A számítógép megjelenésével lehetővé vált, hogy globális változók segítségével adjuk meg a hangok karakterét, melyekből az egyes szemcsékre vonatkozó paramétereket az algoritmus határozza meg. A globális változók (pl. maximális szemcsetranszpozíció) jól kijelölik a percepció alapú tulajdonságokat is. Az additív szintézishez képest – ahol szintén kezelhetetlennek tűnő adatmennyiséggel dolgozunk – ebből a szempontból könnyebb a helyzetünk.

5.2. Megvalósítás

Az eljárást a `HApp_05_01` `(OSX)` `(WIN)` példa illusztrálja, mely legfeljebb nyolc párhuzamos szólamban képes a szemcsék lejátszására, és előre felvett hangfájlokat alkalmaz a szemcsék kitöltésére. A kezelőfelületet az 5.2. ábra mutatja be.

Az interfész felső blokkjában ki lehet választani, hogy melyik hangfájlból és milyen hangerővel válasszon szemcséket a szintetizátor. A középső blokkban a lejátszás módját (folyamatos, vagy amplitúdó-burkológörbével meghatározott), illetve a szemcsék térbeli eloszlását lehet szabályozni.

A szemcsék kiolvasása kétféle módon történhet. „*Automatikus*” esetben a „*lejátszás sebessége*” paraméter állítja be, hogy milyen gyorsan haladjunk végig az eredeti mintán. A „*kézi beállítás*” választásakor megjelenő grafikus vezérlőn („szemcse pozíciója”) jelölhető ki, hogy melyik szakaszát használjuk a hangfájlnak. A szemcséket a program véletlenszerűen választja az így megadott tartományból.

A legalsó blokkban a szemcsehossz (amely egyben meghatározza a követési távolságot is), valamint a transzpozíció paramétereit lehet beállítani.

A program tartalmaz egy kétdimenziós felületet (lásd 5.3. ábra), melyen rajzolással lehet az időben változó paramétereket vezérelni, az egyes frázisokat rögzíteni és visszajátszani két hangszíndimenzió mentén. A felületen állítható a paraméter típusa, valamint minimális és maximális szélsőértéke az x és az y tengelyeken.

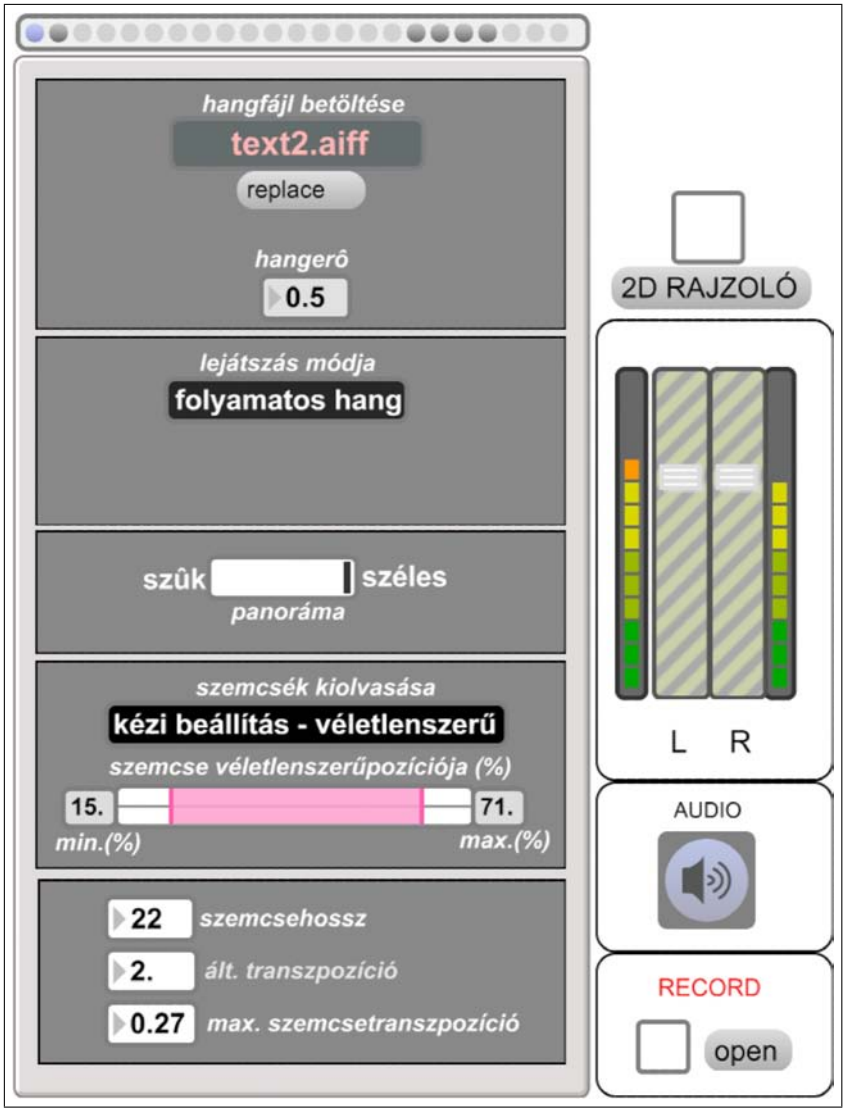
Az 5.1. program egy szinuszos hullámformákat használó granuláris szintetizátor implementálása CSound nyelven (a lekevert hangot az 5.12. hangpélda tartalmazza).

5.3. Összegzés

Hangszín. Mivel elvileg végtelen számú és karakterű spektrumot lehet használni kiindulásul, az eredményül kapott hangzások is elvileg végtelenül nagy hangszínteret fed-

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



5.2. ábra. Egy granuláris szintetizátor kezelőfelülete. A program a szemcséket egy hang leból vágja ki.

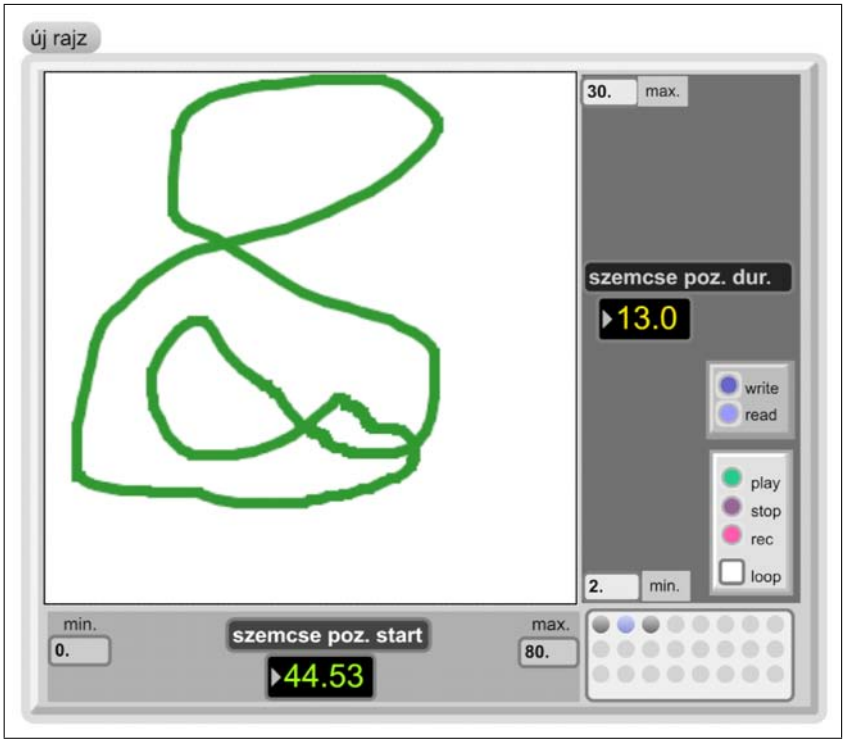
nek le. A hangok minősége nagyban függ a kiindulási jelek milyenségétől. A szintézistechnikára jellemző egyfajta remegő, gyorsan villódzó, érdes hangzás, melynek alapján könnyen azonosítható. A granulálás paraméterei hangsúlyoznak egyes érzékelési jellegeket, mint pl. az érdesség-simaság, a konstans-kaotikus, zajos-harmonikus dimenziókat. Felületes paraméter-beállítások esetén előfordul, hogy sokszor hallott hangzásklisékhez jutunk alkalmazásával.

Paraméterek. A granulálás paramétereinek befolyása a hangzásra könnyen megérthető, jól tanulható. Egyes paraméterek (pl. szemcse hosszúsága, szemcsék követési távolsága) segítségével, megfelelő áttérképezéssel közvetlenül alkothatunk hangszínskálákat.

Műveletigény. A granuláris szintézis nagyon sok szemcsét és paramétert kezel egyszerre, ezért erősen erőforrásigényes. A szólamok és a vezérlő jelek számának csökkentésével lehet optimalizálni a számítógép kihasználtságát.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



5.3. ábra. Gra kus adatbevitel. A két tengelyen a két szabályozni kívánt paraméter szerepel, melyeket legördülő listákból választhatunk ki. A rajzolt gesztust felvehetjük, illetve visszajátszhatjuk.

5.4. Feladatok

1. A legtisztábban úgy érthetjük meg a granuláris szintézis egyes paramétereinek a hatását, ha forráshangként egy egyszerű szinuszhullámot használunk (más esetben nem tudhatjuk, hogy a végeredmény mekkora részben származik a granulálásból és mekkora részben tudható be a forrásjel spektrumának). Hallgassuk meg a granulálás „természetét". Vizsgáljuk meg, hogy az egyszerű szinuszhullámra milyen hatást gyakorol a granuláris szintézis. Először állítsuk be úgy a paramétereket, hogy a végeredmény megegyezzen az eredeti szinuszhullámmal, majd változtassuk egyenként a lehetséges beállításokat. Figyeljük meg, hogy az egyes paraméterek mentén hogyan változik a hangszín!

2. Bizonyos esetekben a granuláris szintézis eredménye függetlenné válik a felhasznált hangforrástól. Keressünk olyan paraméter-csoportokat, amikre ez teljesül!

3. A `HApp_05_01` `(OSX)` `(WIN)` rajzolópanelének segítségével vegyél fel egy 2D-mozgásfolyamatot (*trajektóriát*). Próbáld ki legalább 3 eltérő paraméterpárral a készített trajektória hatásait! Változtasd az egyes paraméterek lehetséges szélsőértékeit is. Hogyan befolyásolják a szélsőértékek a kapott végeredményt?

4. Hozz létre egy tetszőleges hangot a `HApp_05_01` `(OSX)` `(WIN)` programban, majd válassz ki egy paraméterpárt a rajzolópanelen és vegyél fel egy trajektóriát! Kapcsold be a rajzolópanel jobb alsó sarkában található „*loop*" kapcsolót, majd indítsd el a trajek-

Open Access 2026-11-02

tória lejátszását! Hogyan változik a hatás, ha ugyanazt a loopot különböző forráshangokra alkalmazod? Készíts minél több és változatosabb loopot!

5.5. Függelék

5.1. CSound-forráskód. hd_05_01_csound.csd

```
<CsoundSynthesizer>
<CsInstruments>

sr     = 44100
kr     = 4410
ksmps  = 10
nchnls = 1

instr 1
; Granulálás szinuszhullámból
        kpitch     linseg     p5, p3/2, p6, p3/2, p5 ; szemcse-
           hangmagasság burkoló
        kdens      line       p10, p3, p11           ; sürüség burkoló
        krandamp   line       p12, p3, p13           ; amplitudó vé
           letlenszerüség burkoló
        krandpitch expon      p14, p3, p15           ; hangmagasság vé
           letlenszerüség burkoló
        kranddur   expon      p16, p3, p17           ; hossz vé
           letlenszerüség burkoló
        a1         grain      p4, kpitch, kdens, krandamp, krandpitch,
            kranddur, 1, p9, 1
        a2         linen      a1, p7, p3, p8
        out        a2
endin

</CsInstruments>
<CsScore>

        f1 0 4096 10 1
        f3 0 4097 20 2 1

        ; p4=amp  p5=freku1  p6=freku2  p7=felfutás  p8=lecsengés  p9=
           szemcseburkoló-hullámforma_táblázat  p10=dens1  p11=dens2
            p12=amp_off1  p13=amp_off2  p14=pitchoffs1(Hz)  p15=
           pitchoffs2(Hz)  p16=graindur1(max 1)  p17=graindur2(max 1)

        ; p1 p2 p3   p4   p5  p6 p7 p8 p9 p10 p11  p12  p13 p14   p15
           p16 p17

        i1    0  2 1000  440 440  1 .1  3  50  50   50   50   1     1
           .1  .1
        i1    +  2 1000  440 440  1 .1  3   5   5   50   50   1     1
           .1  .1
        i1    +  2 1000  440 440  1 .1  3   5   5 5500 5500   1     1
           .1  .1
        i1    +  2 1000  440 440  1 .1  3   5   5 5500 5500 500   500
           .1  .1
```

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02


```
35          i1     +  2 1000   220 220   1 .1   3    5    5 5500 5500 500    500
                .1   .1
36          i1     +  2 1000   440 440   1 .1   3 200 200 5500 5500 500    500
                .01  .01
37          i1     + 15  600   440 440   1 .1   3    5 200 1000 5500    1      1
                1    .005
38          i1     + 10 1500 1760  60   5 .1   3    5 200  500 1000   10 10000
                1    .01
39          e
40
41  </CsScore>
42  </CsoundSynthesizer>
```

© *Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola*

Open Access 2026-11-02

6. fejezet

Fizikai modellezés

A legtöbb szintézismódszer az akusztikus jelre, a létrehozandó hullámra (illetve annak hangszínére) koncentrál. Legyen szó az additív, a szubtraktív, a granuláris stb. módszerről, a kiindulási pont és az elérendő cél minden egyes esetben a jelfeldolgozás nyelvén fogalmazódik meg. Vegyük észre, hogy a „valódi" (mechanikus) hangszerek tervezése máshogy zajlik: a hangszerkészítő nem egy adott hangszínhez készít egy „mechanikai implementációt", hanem különböző anyagok megformázásával, összekapcsolásával ér el adott hangszíneket. A hang, ami így keletkezik, egyenes következménye a hangszer mechanikai struktúrájának, az azt felépítő anyagok (és a levegő) fizikai tulajdonságainak.

A fizikai modellezés során hasonlóan járunk el: valódi (vagy elképzelt) anyagok fizikai tulajdonságainak szimulálásán keresztül jutunk el a létrehozandó hullámformához.

6.1. Elmélet

A fizikai modellek egy adott fizikai struktúra viselkedését modellezik. Minden egyes modell létrehozása során először felírjuk a vizsgált fizikai rendszer (pl. hangszer) viselkedését leíró egyenleteket. Általános esetben egy ún. *nemlineáris parciális differenciálegyenlet-rendszert* kapunk, aminek az analitikus megoldása[1] általában lehetetlen. Ezért többféle egyszerűsítéssel szokás élni, melyek során a teljes rendszert olyan alrendszerekre bontjuk, amelyek viselkedése (numerikus úton) meghatározható.

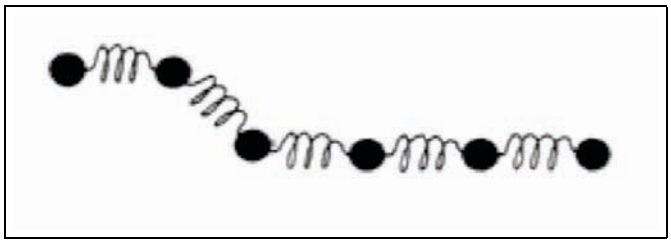
Az egyik megoldás (melyet *tömegpont-rugó modellnek* hívunk), hogy a vizsgált struktúrára úgy tekintünk, mint tömegpontok (tömeggel rendelkező, de végtelenül kicsi részecskék) halmazára, amelyeket apró „rugók" kötnek össze egymással. E „rugók" testesítik meg az egyes tömegpontok közötti kölcsönhatást. A szimuláció során e részecskék rezgéseit számítjuk ki. A folyamatot a 6.1. ábra mutatja be. A keletkező hangot akár egy adott tömegpont, akár a részecskék egy csoportjának összegzett rezgéseiként értelmezhetjük. E módszer, habár általánosan alkalmazható, nagyon számításigényes, ezért valós időben nem, vagy csak nagyon korlátozott mértékben használható.

Az előbbinél egy fokkal egyszerűbb másik megoldás a *modális szintézis*, mely során a fentebb leírt tömegpont-rugó modellt egymástól többé-kevésbé független alrendszerekre (pl.

[1] Analitikus megoldás alatt azt értjük, amikor konkrét (zárt, vagyis további integrálokat már nem tartalmazó) függvényeket kapunk eredményként, így ezen megoldás-függvények ismeretében bármilyen bemenő paraméterre „automatikusan" meg tudnánk határozni a kialakuló hullámformát.

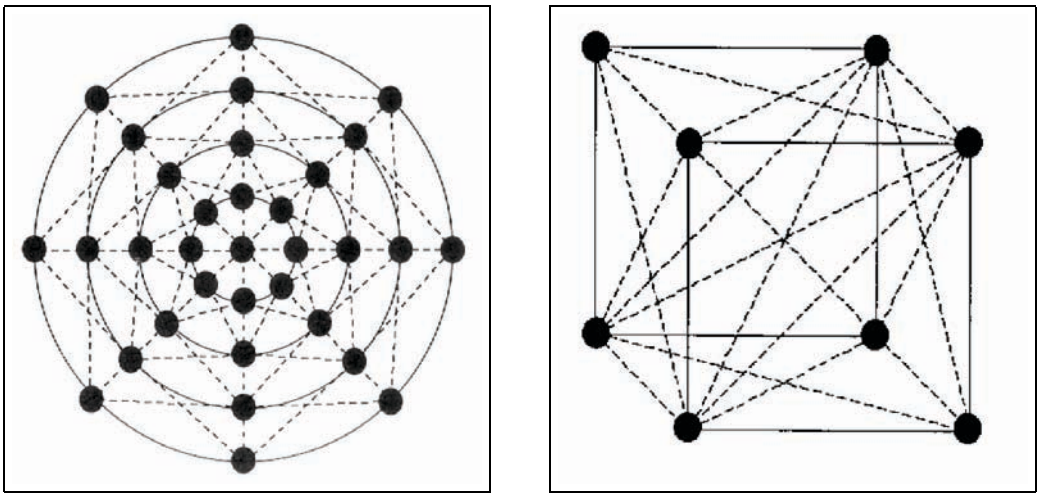
© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



6.1. ábra. Tömegpont-rugó modell. Az ábrán a bal oldali tömegpont kitérített állapotban van, a modellezés során lépésenként modellezzük, hogy e kezdeti kitérés hatására hogyan mozog a teljes rendszer.

hegedű esetén lábra, húrra, hangszertestre stb.) bontjuk, majd a kapott alrendszerek kollektív (vagyis az egész testre kiterjedő, egységes) rezgéseit határozzuk meg. A 6.2. ábra ilyen alrendszereket mutat be. E kollektív rezgések a test *módusai*, az eljárás innen kapta a nevét. A módusok kiszámítása után a következő lépés az illesztések mentén kialakuló *csatolások* vizsgálata. E csatolások határozzák meg, hogy az egyes alrendszerek lehetséges módusai miként vesznek részt a teljes rendszer (a hangszer) rezgéseiben. A modális módszer ilyen értelemben analóg a vizuális programozásban (pl. Max) megismert alapelvvel: az egyes alrendszerek viselkedését „fekete dobozként” képzeljük el, vagyis (a módusok meghatározása után) nem érdekes, hogy ezek az alrendszerek tulajdonképpen hogy működnek – mint ahogy a Max programozásakor sem lényeges az egyes objektumok belső implementációja. Az alrendszerek közti csatolás hasonló szerepet tölt be, mint az egyes Max-objektumok közötti összeköttetések.



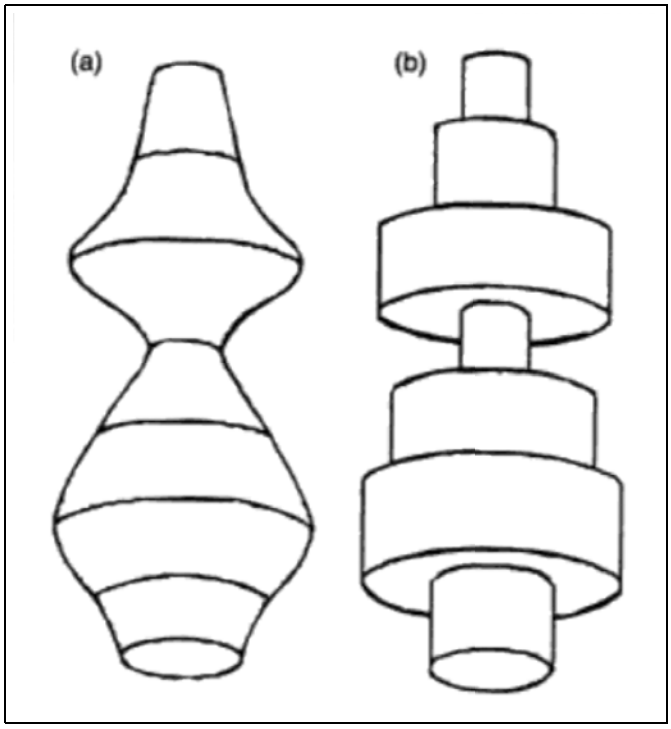
6.2. ábra. Körszimmetrikus (bal) és hasábos (jobb) alrendszerek. A modális szintézis során ilyen alrendszerek kollektív rezgéseit vizsgáljuk.

Egy harmadik módszer *hullámvezetőkre* bontja a vizsgált rendszert. A hullámvezető egy olyan test (pl. cső, húr, membrán), amely a hanghullámok nagyjából veszteségmentes terjedését teszi lehetővé. A 6.3. ábra egy komplex test hullámvezetőkre bontását mutatja. E hullámvezetőknek különböző idejű késleltetéseket feleltetünk meg. A hullámvezetők találkozásánál az egyes hullámok (általában frekvenciafüggő mértékben) veszítenek energiáikból, amit szűrők közbeiktatásával tudunk modellezni. A módszer jelentősége az előző

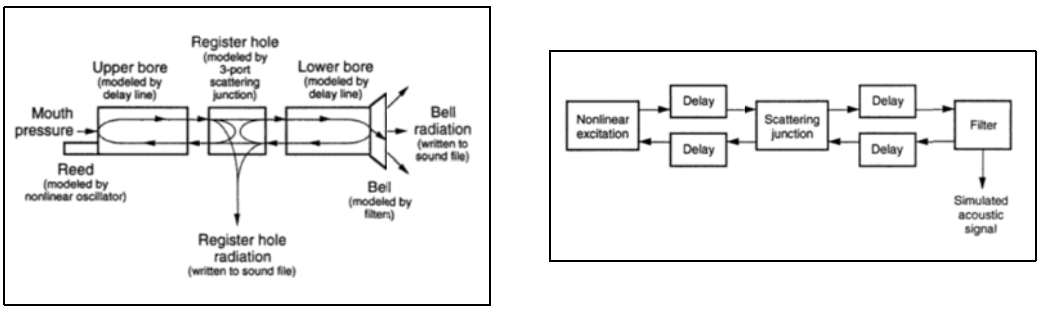
© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

kettővel szemben, hogy a szűrés és a késleltetés jóval kevesebb számítást igényel, mint a közvetlen modellezés. A 6.4. ábra egy fúvós hangszer nagyon egyszerű modelljét ábrázolja.



6.3. ábra. Egy szabálytalan cső közelítése hengeres hullámvezetőkkel: az eredeti, komplex cső (a) és a hullámvezető-modell (b).



6.4. ábra. Egy fúvós hangszer (klarinét) nagyon egyszerű modellje. A fúvókának egy zajszerű gerjesztés, a csőnek különböző késleltetők, míg a hanglyuknak és a tölcsérnek egy keverő és egy szűrő felel meg.

Az eddig leírtak olybá tűnhetnek, hogy fizikai modellezéssel egy adott test *bármilyen* mozgása modellezhető. Habár ez természetesen igaz, a szintézistechnikának ez nem célja.[2] Ezért fontos egyszerűsítésre ad lehetőséget minden egyes fizikai modell elkészítése során, ha feltételezzük, hogy az adott (hangszer)test csak bizonyos típusú *gerjesztésekre* (pl. levegő befúvása egy csőbe, vagy egy húr adott ponton történő megütése) reagál, így számításainkat leszűkíthetjük e gerjesztésekre. Valóban, az eljárásnak nagyon fontos része annak pontos

[2]Például hiába tudjuk meghatározni, hogy egy zongora tömegpontjai hogyan mozognak, miközben arrébb gurítjuk a hangszert, zenei szempontból ez nem érdekes

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

meghatározása, hogy a testet mely pontjaiban lehetséges gerjeszteni, és milyen típusú gerjesztés az, amit e pont(ok)ban várunk. Emiatt külön (fizikai) modellt szokás alkotni az egyes gerjesztéstípusokra is (pl. vonó és húr érintkezése). A csatolás jellege, amin keresztül e gerjesztések rezgésbe hozzák a hangszertestet, központi szerepet játszik a hang végső kiszámításában.

6.2. Megvalósítás

A továbbiakban egy korai, ám nagyon sikeres eljárást, a megpendített húrokat utánzó *Karplus–Strong algoritmust* ismertetjük. A modell egy húrból (késleltető) és annak egy csatlakozási pontjából (szűrő) indul ki. A késleltetési idő a húr hosszának (és ezen keresztül a hangmagasságnak) felel meg, míg a szűrő karakterisztikája azt határozza meg, hogy a húrban terjedő hullám melyik frekvenciájából mennyit veszít, amikor a csatlakozási ponton visszaverődik (ez végső soron a hangszínt befolyásolja). Ez a visszaverődés a modell nyelvére lefordítva azt jelenti, hogy a szűrt jelet visszavezetjük a késleltető bemenetére. A húr pengetését úgy modellezzük, hogy kiindulási állapotként (ez a „megpendítéskor” a késleltetőben tárolt hullámforma) fehérzajt veszünk. A 6.1. program mutatja az eljárás implementációját CSound nyelven (az eredményt a 6.1. hangpélda tartalmazza).

A Max implementációt a 6.5. ábra mutatja. Itt a szűrést egy változtatható frekvenciájú aluláteresztő szűrő végzi, melynek pillanatnyi frekvenciáját a jobb oldali burkológörbeszerkesztő határozza meg (a burkológörbe minden hangindításkor újraindul). A „*zajimpulzus LP frekvencia*” a késleltető kezdeti feltöltésére használt fehérzaj szűrését teszi lehetővé, míg a „*feedback*” beállítás a visszacsatolás erősségét adja meg. A „*torzítás mértéke*” segítségével a kimeneti jelet túlvezérelhetjük (lásd a 13. fejezetet). Ez természetesen nem közvetlen része az algoritmusnak.

Figyeljük meg, hogy a hang fényessége erősen függ a kezdeti zaj spektrumától: fehérzaj esetén fényes, míg erősen szűrt kezdeti jel esetén sötét hangszínt kapunk. A visszavert hullámot moduláló szűrő karakterisztikája elsősorban a hang öblösségét, rezonáns mivoltát szabályozza, de közvetlenül hatással van a hang időtartamára is (a továbbfejlesztett modellekben ezt a két tulajdonságot általában külön paraméterek szabályozzák).

A modell általánosítható dobokra, vonóval gerjesztett húrokra, de akár még egyes fúvós hangszerekre is.

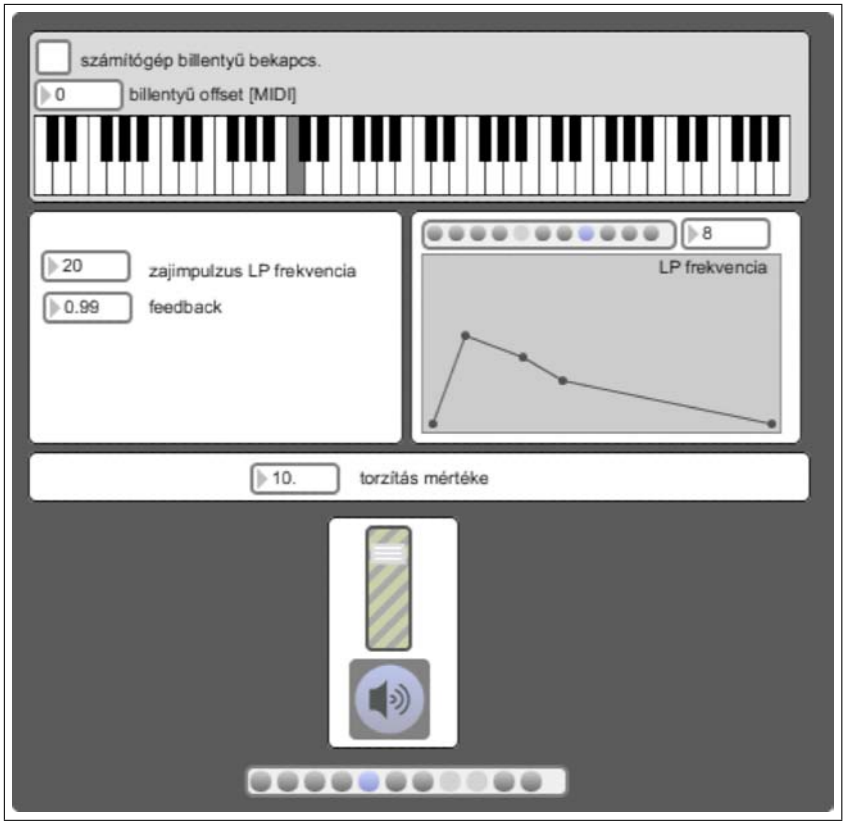
A Karplus–Strong algoritmus, habár nagyon egyszerű és hatásos, csak alapjaiban mutatja be a technikában rejlő lehetőségeket. A 6.6. ábra egy komplex rézfúvós modell vezérlőpanelét mutatja. A `http://www.youtube.com/watch?v=vyZAliNfwUk` címen pedig egy fizikai modellezésre épülő szintetizátor-család, a Yamaha VL sorozat egyes tagjait hallhatjuk (pl. sakuhacsi, jazz-trombita, hegedű stb.).

6.3. Összegzés

Hangszín. A fizikai modellezés – főleg bonyolultabb rendszerek esetén – igen realisztikus hangszínt hoz létre, bizonyos esetekben szinte tökéletesen reprodukálja az eredeti hangszer hangzását. Ám mivel a paraméterek és a hangszín között általában nincs közvetlenül meghatározható kapcsolat, ezért nagyon nehéz a hangszínt előre tervezni.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



6.5. ábra. Karplus–Strong algoritmus Max programnyelven. A példát a `HApp_06_01` `(OSX)` `(WIN)` *program tartalmazza.*

Ez fokozottan igaz, ha nem létező hangszerekből indulunk ki, vagy amikor fizikailag irreális értékeket adunk a paramétereknek.

Paraméterek. A módszer általában a modellezni kívánt rendszer főbb fizikai paramétereit használja, így akkor válik intuitívvá, ha az előadóművész mozgását rögzítő szenzorokkal (pl. egy adott hangszert modellező kontrollerrel) vezéreljük őket. A paraméterek manuális megadása sokszor nem, vagy nehezen lehetséges, hiszen (modelltől függően) nagyon sok adatra lehet szükség, melyek nem kapcsolódnak közvetlenül a kapott hangzáshoz.

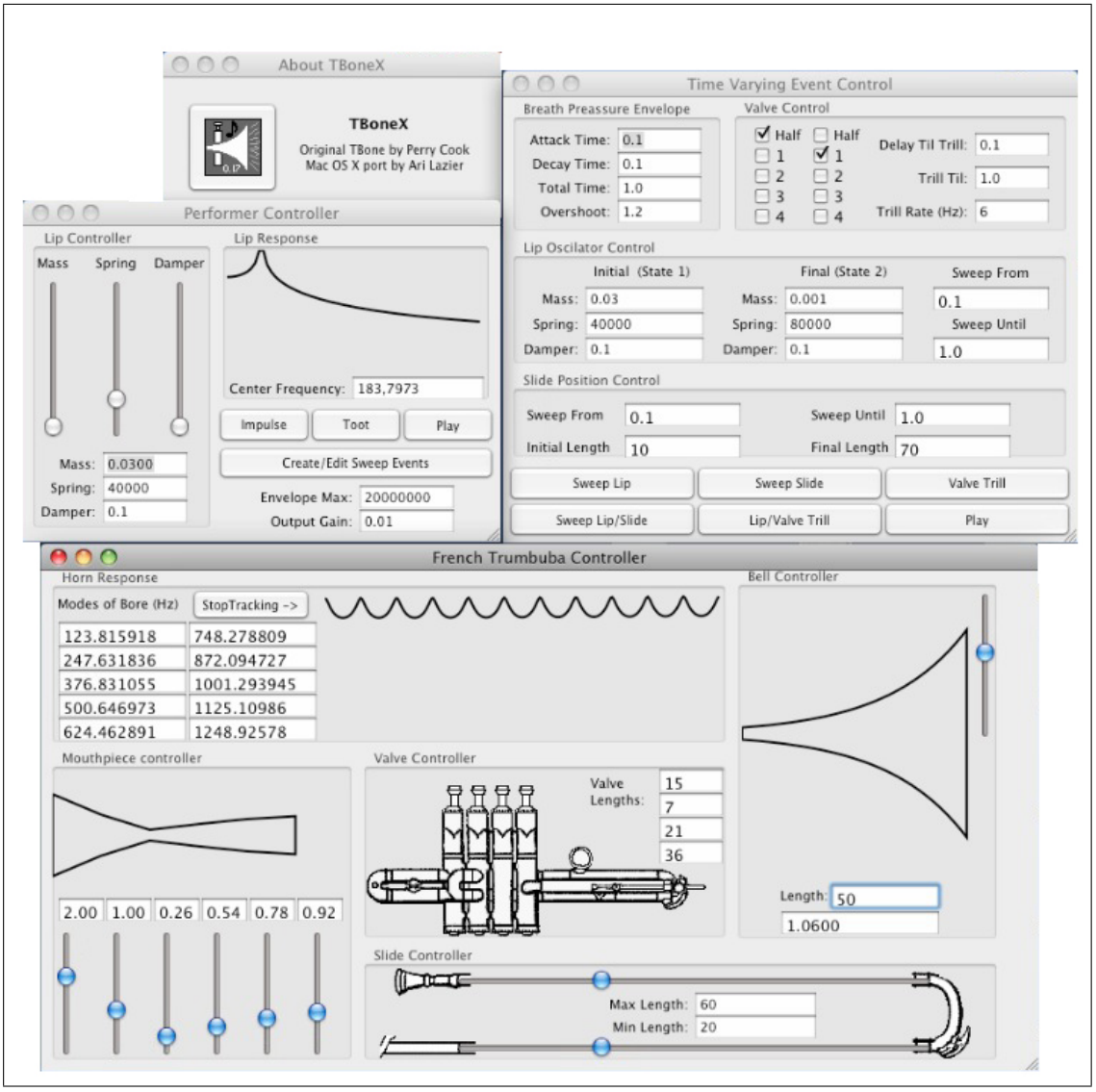
Műveletigény. Mivel a kiindulási rendszerek és az alkalmazott erőforrások függvényében teljesen eltérő algoritmusokat kaphatunk, a skála igen nagy. Bizonyos algoritmusok (pl. Karplus–Strong) meglepően kevés műveletet igényelnek, míg mások erőforrásigénye akkora lehet, hogy a valós idejű szimulációt lehetetlenné teszik.

6.4. Feladatok

1. A `HApp_06_01` `(OSX)` `(WIN)` segítségével hozd létre a 6.2., a 6.3. és a 6.4. hangpéldákat!

2. Ahogy említettük, a Karplus–Strong algoritmus korai formájában a szűrő a fényességre és az időtartamra egyaránt hatással van. Vizsgáld meg ezt a hatást! Készíts hangokat

Open Access 2026-11-02



6.6. ábra. A CCRMA által fejlesztett TBone kezelőfelülete. A szoftver rézfúvós hangszerek (elsősorban harsona és trombita) modellezésére alkalmas.

azonos egyéb beállításokkal (pl. hangmagasság, visszacsatolás), de eltérő szűrésekkel – konstans burkológörbét választva! Hogyan módosítja az egyes hangokat, ha a burkológörbe nem konstans?

3. A visszacsatolás mértékét változtatva keressük meg, hogy milyen tartományban viselkedik a modell húrként és mikor hasonlít jobban egy rezgő membránra (pl. dob)! Hogyan történik a két hangszín között az átmenet?

4. Befolyásolja a zajos (dob jellegű) hangszíneket a szűrő? Ha igen, hogyan?

6.5. Függelék

6.1. CSound-forráskód. `hd_06_01_csound.csd`

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02


```
<CsoundSynthesizer>
<CsInstruments>

sr     = 44100
kr     = 4410
ksmps  = 10
nchnls = 1

instr 1
; Karplus Strong Algoritmus

        icps        =           cpspch(p5) ; hangmagasság
        asig        init        0          ; asig varváltozó bevezetése
        kcount      init        1/icps*kr  ; loop számláló (lásd később)

        ; 1/cps hosszúságú késleltetés:
        adel        delayr      1/icps     ; 1/cps sec késleltetés
        asig        tone        adel,sr/2  ; kimenet szürése

        ; késleltetö lánc feltöltése zajjal:
        if (kcount < 0) kgoto continue   ; átugorni, ha megtelt
kloop:
        asig        rand        p4,-1
        kcount      =           kcount - 1
continue:
        delayw      asig                   ; késleltetö bemenete
        if (kcount >= 0) kgoto kloop
        out         asig
endin

</CsInstruments>
<CsScore>

        ; instr start hossz   amp magasság
        i01          0     2 10000     6.00
        i01          +     .  5000     6.00
        i01          +     . 10000     7.00
        i01          +     .     .     8.00
        i01          +     .     .     9.00
        e

</CsScore>
</CsoundSynthesizer>
```

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

7. fejezet

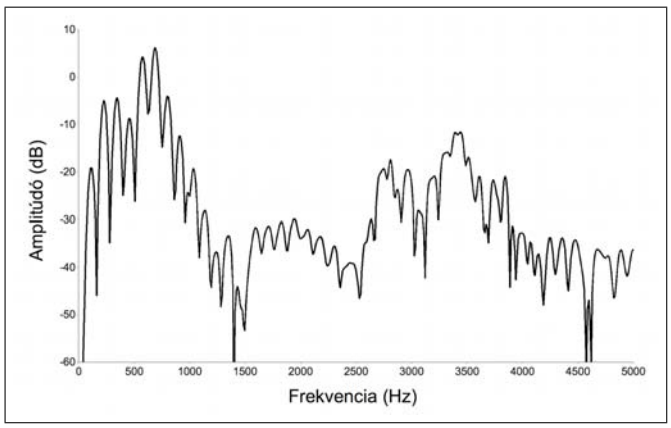
Formáns szintézis

Az itt bemutatott technika akár az additív (3. fejezet), akár a szubtraktív (4. fejezet) szintézis, de még a fizikai modellezés (6. fejezet) szemszögéből is értelmezhető,[1] így – technikai értelemben – semmiképp sem tekinthető önálló szintézismódszernek. Ám pont amiatt, hogy e három eltérő eljárás metszetében található, nagyon karakteres és jellemző a hangszíntér, ami létrehozható vele. Így a módszerre *zenei okokból* tekintünk úgy, mint önálló szintézistechnikára. Az eljárással elsősorban időben statikus hangszerhangokat, illetve magánhangzókat szintetizálhatunk.

7.1. Elmélet

7.1.1. A formánsok

Ha megvizsgáljuk egy valódi hangforrásból származó, kitartott hang spektrumát, látni fogjuk, hogy az egyes csúcsok általában nem élesek. Olyan területeket fedezhetünk fel, amelyek kiemelkednek ugyan a háttérzajból, de egy élesen meghatározható frekvencia helyett a spektrum egy nagyobb – akár több száz Hz szélességű – területét foglalják el (lásd a 7.1. ábrát).



7.1. ábra. Egy „a” magánhangzó spektrumának alsó tartománya. Figyeljük meg, ahogy a legjellemzőbb formánsok kiemelkednek a spektrumból.

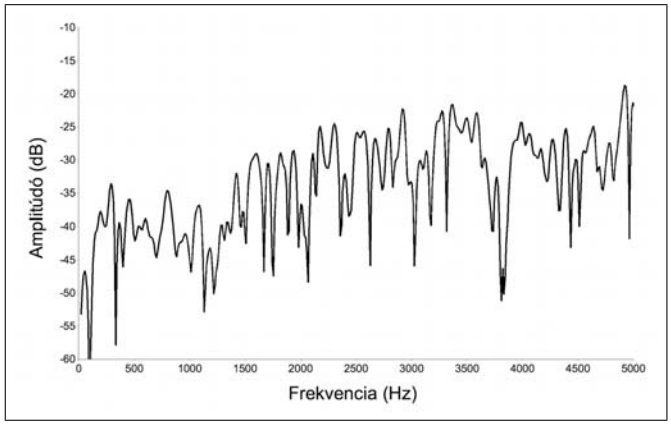
[1] Általában e legutóbbihoz szokás sorolni.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

Ezeket a régiókat nevezzük *formánsoknak*. Jellemzésükhöz három adat egyidejű megadása szükséges: a *sávközép-frekvencia*, a *sávszélesség* és a *hangerő*. Ez megegyezik azzal a három adattal, amivel a 4. fejezetben a sáváteresztő szűrőket jellemezhettük. A kapcsolat nem véletlen: látni fogjuk, hogy a formánsok előállításának egyik módja szűréssel történik.

Erősen zajos, vagy időben gyorsan változó hangoknál nem szokás formánsokról beszélni. Ennek oka, hogy a paraméterek, elsősorban a sávközepek és szávszélességek (zajos hangoknál, lásd a 7.2. ábrát), illetve ezek „élettartama" (gyorsan változó hangoknál) nehezen, vagy egyáltalán nem határozhatók meg. Emiatt a formánsok fogalmát elsősorban a már említett, időben lassan változó, nem túlságosan zajterhelt hangok leírására alkalmazzuk. Ilyen a legtöbb fúvós hangszer hangja, illetve az emberi magán- és zöngés mássalhangzók.



7.2. ábra. Egy „sz" mássalhangzó spektrumának alsó tartománya. Láthatjuk, hogy ez a spektrum nem jellemezhető pusztán néhány csúcs megadásával.

A formánsokon alapuló szintézist először a korai beszédszintetizátorok alkalmazták (leginkább a magánhangzók előállítására), ezért először dióhéjban áttekintjük az emberi hangképzést, ezen belül is a magánhangzóképzés egyik korai modelljét. Ezek után bemutatjuk azt a két legelterjedtebb módszert, amit formánsok előállítására használnak. Megjegyezzük, hogy az emberi magánhangzók helyett kiindulhatnánk egy tetszőleges fúvós hangszer modelljéből is. A választásunk leginkább történeti okokból esett a magánhangzó-képzésre, de látni fogjuk, hogy maga a módszer ugyanúgy alkalmazható hangszerek hangjaira is.

7.1.2. A magánhangzók képzése

A köznapi beszéd forrása a *tüdőkből* kiáramló levegő, melyet a fölöttük elhelyezkedő *gége* és *vokális traktus* együttese alakít beszédhangokká. A hangképzésben a gégében elhelyezkedő *hangszálak* és a *hangrés*, a vokális traktus üregei (a legfontosabbak a *garatüreg*, a *szájüreg* és az *orrüreg*), valamint a *nyelv*, a *fogak*, az *ajkak*, a *szájpadlás* és néhány további szerv vesz részt.

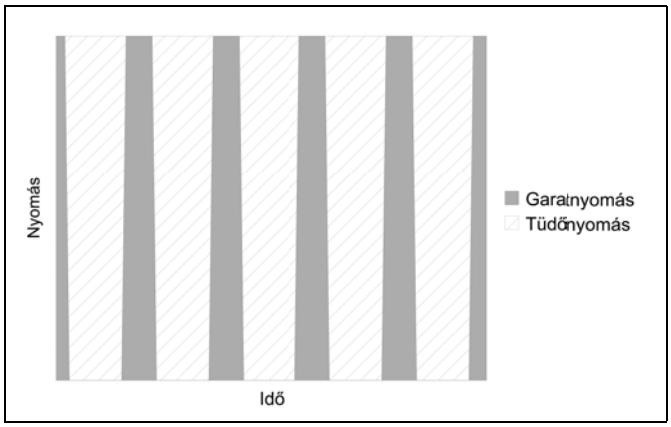
A legegyszerűbb beszédhangok a *magánhangzók*, melyek képzésekor a tüdőkből kiáramló levegő rezgésbe hozza a hangszálakat, majd a teljesen nyitott szájüregen keresztül távozik, további akadályokba nem ütközve. Ezzel szemben a *mássalhangzók* létrejöttekor a fentebbi szervek közül legalább egy úgy helyezkedik el, hogy akadályt képezzen a levegő útjában.

A fenti megfigyelés alapján nagyon egyszerűen modellezhető a magánhangzók létrejötte.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

A hangszálak, illetve a hangrés szerepe, hogy azok a tüdőkből egyenletes sebességgel kiáramló levegőt szabályos időközönként elzárják, periodikusan megszakítva ezzel a folyamatos légáramlást. A hangnyomást, amit a képzeletbeli hangforrásunk létrehoz, a 7.3. ábra illusztrálja. Habár a valóságban a garaton átáramló levegő nyomása ennél jóval bonyolultabb képet mutat,[2] látni fogjuk, hogy ezzel az igen egyszerű közelítéssel is meglepően jó eredményeket lehet elérni.



7.3. ábra. A tüdőkből kiáramló levegő, illetve a garaton (hangszálak, hangrés) áthaladó légoszlop nyomásának vázlatos képe az idő függvényében.

A vokális traktust egy szabálytalan keresztmetszetű, ám az alakját a hangképzés során nem változtató csőnek tekintjük.[3] Ennek modellezése jelenti a nagyobb kihívást. Bizonyítható, hogy egy szabályos, mindkét oldalán nyitott cső a hossza kétszeresének megfelelő hullámhosszú hangot, illetve ennek felhangjait felerősíti. Azt azonban sokkal nehezebb meghatározni, hogy egy olyan bonyolult idom, mint amilyen a vokális traktus, pontosan melyik frekvenciákat emeli ki. Ezért a gyakorlatban ahelyett, hogy becslést próbálnánk adni a vokális traktus pontos alakjára (majd ebből az általa felerősített frekvenciatartományokra), inkább közvetlen mérésekkel szokás meghatározni, hogy egy-egy magánhangzó kiejtésekor milyen formánsok jelentkeznek a spektrumban. Az így megmért adatok átlagolásával olyan formánscsoportokat kapunk, amelyek jól jellemzik a szintetizálni kívánt magánhangzókat. Egy ilyen mérés eredményét mutatja be a 7.1. táblázat és a 7.4. ábra.

7.1. táblázat. Tenor énekesek által énekelt „a” magánhangzó legjellemzőbb formánsai.

Formáns sorszáma	I	II	III	IV	V
Sávközép (Hz)	650	1080	2650	2900	3250
Sávszélesség (Hz)	80	90	120	130	140
Hangerő (dB)	0	−6	−7	−8	−22

A fenti leírásból nyilvánvaló, hogy (legalábbis a magánhangzóképzés során) a formánsok függetlenek az alapfrekvenciától, hiszen míg előbbit egyedül a vokális traktus rezonáns

[2] Elég csak arra gondolnunk, hogy a hangrés összezáródása nem pillanatszerűen történik, vagy arra, hogy amíg a garati rész zárva van, addig a tüdőkből érkező levegő nyomása megnő.

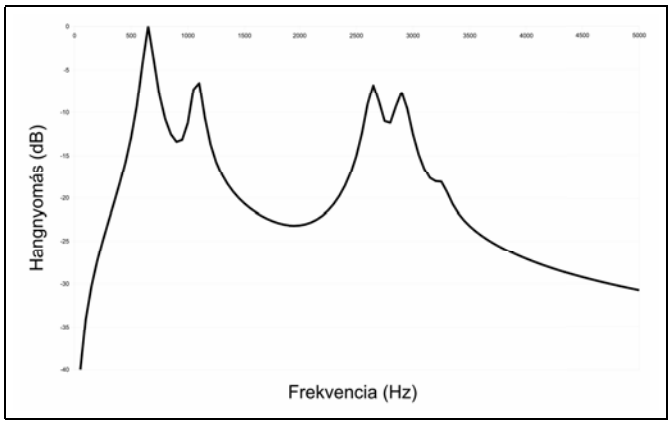
[3] Megjegyezzük, hogy mivel az orrüregen minimális mennyiségű levegő távozik a magánhangzók képzésekor, ezért e képzeletbeli csövet általában elágazásmentesnek feltételezzük.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

tulajdonságai határozzák meg, utóbbi a hangrés által keltett szabályos megszakítások következménye. Az alapfrekvencia és a formánsok függetlensége abból adódik, hogy a hangrés viselkedése gyakorlatilag nem befolyásolja a vokális traktust. Ugyanez elmondható a legtöbb hangszerről is: a formánsokat általában a hangszertest rezonanciái határozzák meg, melyek többnyire függetlenek a hangkeltés módjától (ami pedig az alaphangért felelős).

A formánsok jelentőségének megértéséhez gondoljuk meg, hogy a beszéd során az adott magánhangzó „jelentését” egyedül a formánsok hordozzák; az, hogy egy adott hangzást milyen magánhangzóként értelmezünk, független az adott hang hangmagasságától. Egy adott magánhangzó formánsszerkezetét (és az ebből adódó hangszínt) tudatunk konkrét jelentéssel ruházza fel, mely alapvető szerepet játszik az emberi kommunikációban. Érthető hát, hogy – hasonló módon – az egyes hangszerek formánsai is meghatározó szerepet játszanak az adott hangszer jellegzetes hangszínérzetének kialakításában.



7.4. ábra. Tenor énekesek által énekelt „a” magánhangzó legjellemzőbb formánsai.

7.2. Megvalósítás

7.2.1. Formáns szintézis szűrőkkel

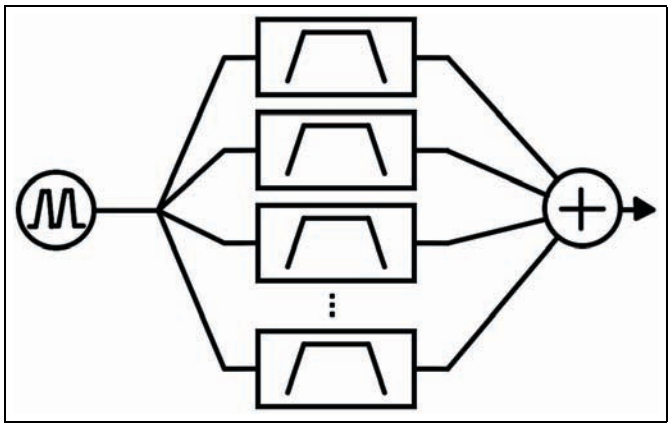
Az imént ismertetett modell közvetlenül megvalósítható egy impulzusforrás megfelelő szűrésével, amint az a 7.5. ábrán látható.

A gége szerepét egy impulzusgenerátor veszi át, melynek két paramétere van: az egyik az *alapfrekvencia*, mely a hangrés zárásának gyakoriságát adja meg, míg a másik az *áteresztési arány*, ami kifejezi, hogy a hangrés két nyitása között mennyi időn keresztül áramolhat ki a levegő (ahol a 0% azt jelenti, hogy a hangrés soha nem nyílik ki, míg a 100% azt, hogy soha nem záródik be[4]). Ennek az impulzusgenerátornak a kimenetét egy ekvalizátorba vezetjük, melyek paramétereit (*hangerők*, *sávközepek*, *jósági tényezők*/*sávszélességek*) az egyes formánsok megfelelő adataiból származtatjuk. A módszert a 7.6. ábrán bemutatott interaktív példa segítségével próbálhatjuk ki.

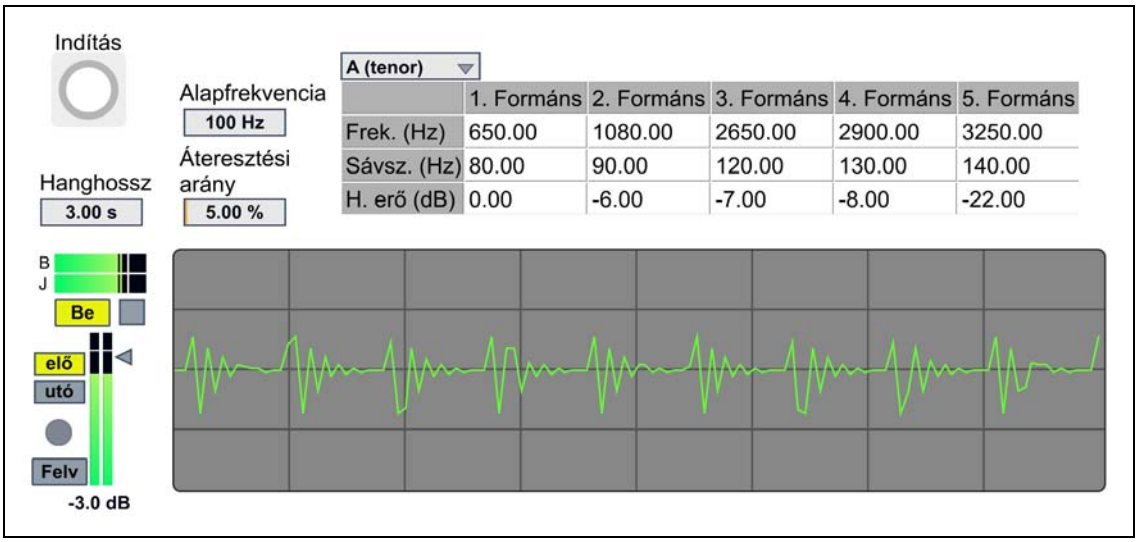
A jobb felső legördülő listából tetszőleges magánhangzót választhatunk. Az egyes szintézis-paraméterek (hanghossz, alapfrekvencia, áteresztési arány), valamint a kimeneti hangerő beállítása után a bal felső gomb megnyomásával hallhatjuk a szintézis eredményét

[4] Az áteresztési arány nem más, mint a 2. fejezetben az impulzussorozat definiálásánál bevezetett δ változó.

Open Access 2026-11-02



7.5. ábra. Ekvalizátorral megvalósított formáns szintézis blokkdiagramja.



7.6. ábra. `HApp_07_01` `(OSX)` `(WIN)`: Formáns szintézis szűrőkkel.

(ezt illusztrálja a 7.1. hangpélda). Az egyes formánsokat a jobb felső sarokban levő táblázat mutatja be, míg a jobb alsó részen levő oszcilloszkóp a szintetizált jel időbeli lefutását mutatja meg. Figyeljük meg, hogy az áteresztési arány változtatása hogyan befolyásolja a hang „életszerűségét”.

7.2.2. Additív formáns szintézis

A szűrésen alapuló módszer – habár digitális megvalósítása felettébb egyszerű – nem túl hatékony. Ennek legfőbb oka, hogy a művelet annyi szűrőt használ, ahány formánsból áll az előállítandó hang. Mivel a jó minőségű digitális szűrők műveletigénye magas, ezért – amennyiben sok formánssal dolgozunk – a módszer nagyon magas számítási kapacitást igényel. Ám a szűrőket egy egyszerű módszerrel száműzhetjük: kiszámíthatjuk a szűrőrendszer *impulzusválaszát*. Ez egy olyan jel, amit akkor kapunk, ha a 7.5. ábra szűrőit egy impulzus-

Open Access 2026-11-02

generátor helyett egy olyan forrásra kapcsoljuk, ami csak egyetlen impulzust ad ki.[5] A módszer lényege, hogy az így kapott jelet eltároljuk egy hullámtáblába; ezzel a formáns szintézis problémáját gyakorlatilag visszavezettük a 2. fejezetben megismert hullámforma-ismétlésre, melynek műveletigénye jóval kisebb, mint a szűrőrendszeré.

Az imént vázolt technikát a szakirodalom FOF-szintézis[6] néven ismeri. A gyakorlati megvalósítás során kihasználhatjuk, hogy a szűrőrendszer impulzusválasza matematikai eszközökkel előre kiszámítható: egy sáváteresztő szűrő a (csak elméletben létező) „végtelenül rövid" impulzust egy szabályos, exponenciálisan lecsengő szinuszhanggá alakítja. Ennek ismeretében a szintézis alapját képező hullámforma a formánsok paraméterei alapján – az egyes szűrők matematikailag meghatározható válaszait összeadva – előre kiszámítható. Megjegyezzük, hogy a gyakorlatban legtöbbször nem a fent említett, tisztán exponenciális lecsengést szokás alkalmazni, hanem ennek valamilyen „finoman" módosított változatát, ami figyelembe veszi, hogy a hangrés nyitott állapotában nem „végtelenül rövid" idő alatt áramlik át a levegő. A módszert a 7.7. és a 7.8. ábrán bemutatott interaktív példák segítségével próbálhatjuk ki.

A 7.7. példaprogramban három különböző módon tudunk formánsokat létrehozni, melyek mindegyikét egy-egy panellel vezérelhetjük. Az egyes panelek a felettük levő feliratokra kattintva aktiválódnak. A bal oldali panel egyetlen formánst generál, melynek megadhatjuk középfrekvenciáját, hangerejét és sávszélességét, valamint a generálás alapfrekvenciáját. A középső panellel három formánst tartalmazó jelet hozhatunk létre, melyek frekvenciája és hangereje önállóan szabályozható. A jobb oldali panel ugyanúgy három formánst hoz létre, ám a kialakuló hangot egy burkológörbével modulálhatjuk, így egy dorombhoz hasonlító hangszínt kapunk. E legszélső panelen a „*véletlen formánsadat generálása*" kapcsolóval minden egyes hanghoz más-más formánsokat gyárthatunk (amint az a 7.2. hangpéldában hallható). Ekkor a „doromb" alaphangja nem változik, ám a változó formánsok miatt úgy halljuk, hogy dallamokat hozunk létre; így működik a valódi doromb is.

A program alsó felében levő kijelzők a FOF és a kimenő hang időbeli alakját és frekvenciáit mutatják. Figyeljük meg a bal felső kijelzőn, hogy az egyes paraméterek változtatásának hatására hogyan változik maga a hullámforma! Láthatjuk, hogy (egyetlen formáns esetén) minél kisebb a sávszélesség, annál jobban fog hasonlítani a FOF egy szinuszhullámra; nagy sávszélességek esetén pedig egy rövid, szinuszos szemcsére. Több formáns esetén a végeredményként kapott bonyolultabb FOF úgy születik, hogy az egyes formánsok által definiált függvényeket összeadjuk. Ezt láthatjuk a 7.7. példa középső panelén, de a 7.8. programban is.

Az utóbbiban alkalmazott formánsfüggvény alakja:

$$f(t) \sim \begin{cases} \frac{1}{2}\Big(1-\cos(\beta t)\Big)\mathrm{e}^{-\alpha t}\sin(\omega t+\phi_0), & t < \frac{\pi}{\beta} \\ \mathrm{e}^{-\alpha t}\sin(\omega t+\phi_0), & \text{egyébként,} \end{cases} \tag{7.1}$$

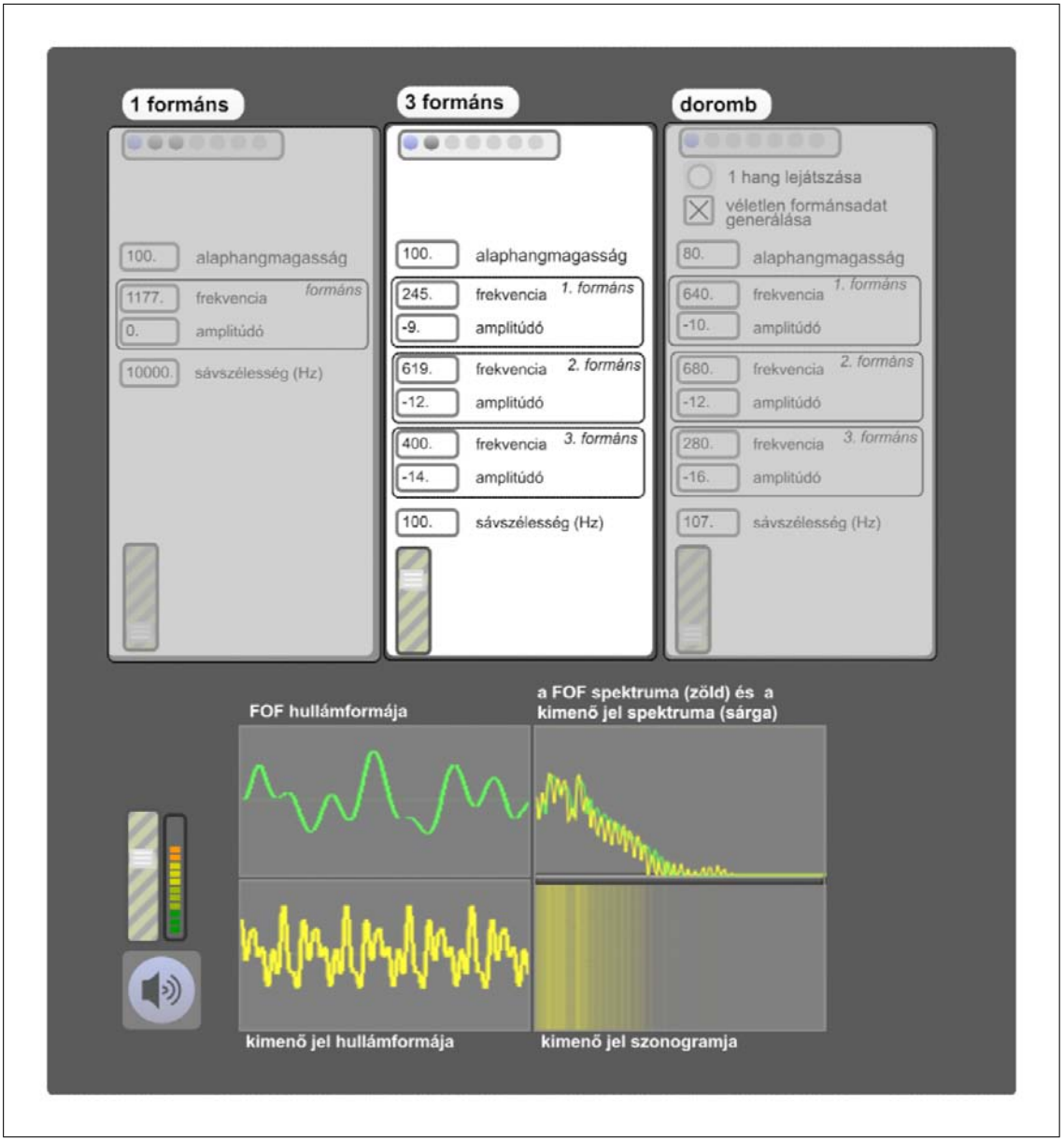
ahol ω a formáns sávközép-frekvenciájából, míg α a sávszélességből származtatható. Az „ideális impulzustól" való eltérés a β paraméteren (*skirt factor*) keresztül jelenik meg: $\beta \to \infty$ esetén kapjuk a „végtelenül rövid" impulzusra kiszámított választ. A függvényt a 7.9. ábra illusztrálja.

[5] Ez annak felel meg, mint amikor egy adott magánhangzót csak egyetlen „pillanatra" ejtünk ki (vagyis, csak egyetlen „kattanást" hozunk létre hangszálainkkal – az így képzett hang egy rövid kattanáshoz, megszakított gurgulázáshoz hasonlít).

[6] A francia *fonction d onde formantique* (formáns hullámfüggvény) kifejezés alapján.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

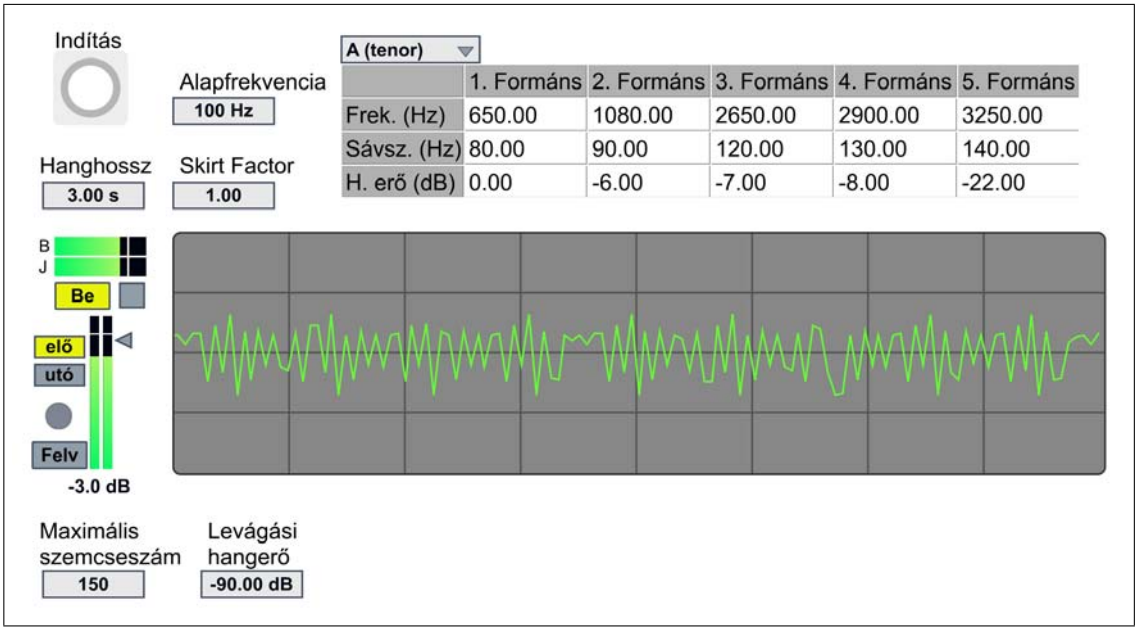


7.7. ábra. `HApp_07_02` `(OSX)` `(WIN)`: általános FOF-szintézis. A három panel eltérő módokon hoz létre formánsokat.

Ha figyelmesebben megvizsgáljuk a (7.1) egyenletet (illetve a 7.9. ábrát), látni fogjuk, hogy a jel amplitúdója, habár idővel igencsak lecsökken, sosem lesz nulla. Ez a köznapi nyelvre lefordítva azt jelenti, hogy az (elméletileg kiszámított) válasz végtelen sokáig tart, vagyis egy végtelen nagyságú hullámformát kellene alkalmaznunk, ami nyilvánvalóan lehetetlen. Ezért egy ponton le kell vágnunk a kiszámított jelet. Minél rövidebb idő után történik ez a vágás, annál pontatlanabb a szintetizált hang, ám minél hosszabb hullámtáblával dolgozunk, annál magasabb a szintetizátor műveletigénye. Utóbbinak az az oka, hogy amennyiben a hullámtábla hosszabb, mint a két impulzus között eltelő idő, úgy oly módon kell belekezdenünk az új hullámtábla-visszajátszásba, hogy az előző impulzusra adott választ még nem

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



7.8. ábra. `HApp_07_03` `(OSX)` `(WIN)`*: FOF-szintézis. Először a jobb felső legördülő listából kiválasztunk egy tetszőleges magánhangzót, majd az egyes szintézis-paraméterek (hanghossz, alapfrekvencia, skirt factor, maximális szemcseszám, levágási hangerő), valamint a kimeneti hangerő beállítása után a bal felső gomb megnyomásával hallhatjuk a szintézis eredményét. Az egyes formánsokat a jobb felső sarokban levő táblázat mutatja be, míg a jobb alsó részen levő oszcilloszkóp a szintetizált jel időbeli lefutását mutatja meg. A 7.3. hangpéldát ezzel a programmal hoztuk létre.*

játszottuk le teljesen.[7] A teljes folyamatot a 7.10. ábra mutatja be.

Hogy a végtelen lecsengést kiküszöböljük, az interaktív példánkban egyrészt megadunk egy *levágási hangerőt* (ez a kiszámított válaszfüggvény legnagyobb és legkisebb amplitúdójának aránya), illetve egy *maximális szemcseszámot* (ez határozza meg, hogy amennyiben az eltárolt hullámforma hosszabb, mint az impulzusok között eltelő idő – vagyis, amennyiben a fent leírt „többszörös kiolvasáshoz” kell folyamodnunk –, úgy egyidejűleg legfeljebb hány példányban történhet meg a hullámtábla kiolvasása).

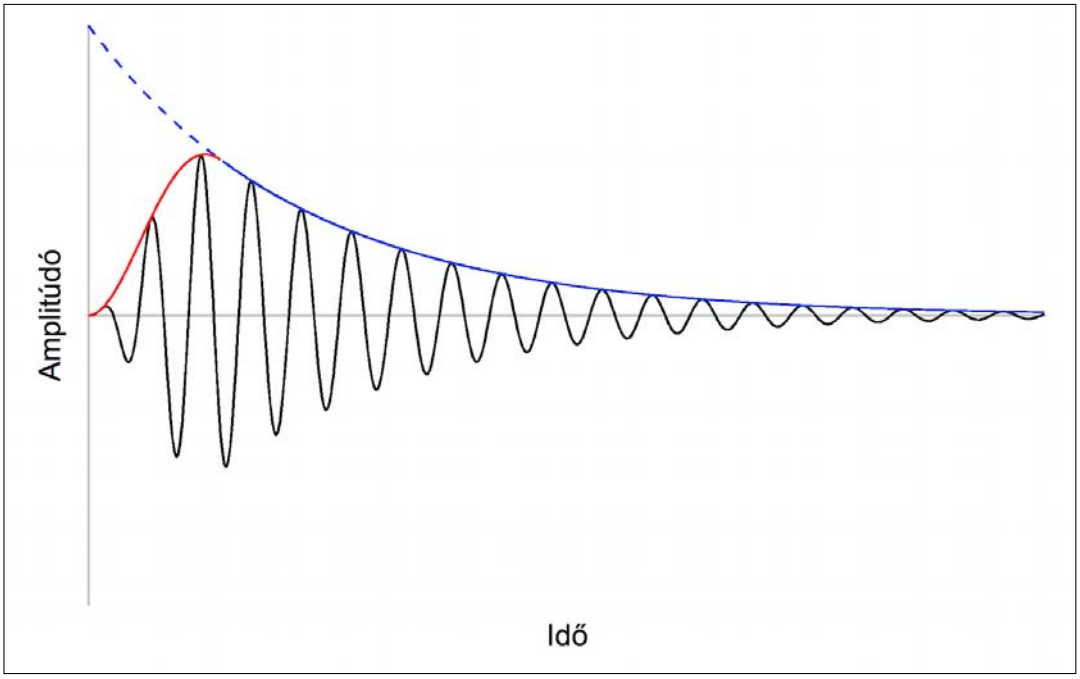
Az eddigi interaktív példák csak statikus hangokat produkáltak. A 7.11. ábrán bemutatott program segítségével kipróbálhatjuk, hogy az egyes hangzókat hogyan tudjuk egymásba fonni. A hangzók számának beállítása után lehetőségünk nyílik arra, hogy az egyes magánhangzókat egyenként válasszuk ki, legördülő listákból. A program kétféle üzemmódot ismer:

Impulzus üzemmódban a hanghossz beállítása után, az indítógomb megnyomásával egy olyan hangot hallhatunk, ahol az egyes magánhangzók egymásba folyva hangzanak el. Ezt hallhatjuk a 7.4. hangpéldában is.

Folytonos üzemmódban egy csúszka segítségével magunk állíthatjuk be, hogy a magán-

[7] Ez az egyetlen eltérés a „klasszikus” hullámforma-ismétléshez képest. Míg a hullámforma-ismétlésnél a hullámtábla időtartama megegyezik a periódusidővel (így nincs szükség az itt leírt „többszörös kiolvasásra”), addig a FOF-szintézis során az eltárolt hullámforma időtartama független a kiolvasás periódusidejétől. Korábban megállapítottuk, hogy a formánsok függetlenek az alaphangtól; ez a viselkedés ennek a megállapításnak az egyenes következménye.

Open Access 2026-11-02



7.9. ábra. A (7.1) egyenlet által leírt függvény. A kék vonal az α által megszabott exponenciális esést, a piros vonal a β által szabályozott koszinuszos indítást mutatja. A két görbe $t = \frac{\pi}{\beta}$-nál találkozik. A végeredményt (a modulált szinuszhullámot) a fekete jel ábrázolja.

hangzók közti átmenetek melyik állomását szeretnénk hallani.

Ez az egyszerű példaprogram jól illusztrálja, hogy két hang formánsainak ismeretében viszonylag egyszerűen hozhatunk létre átmeneteket az eredeti hangjaink között.

7.3. Összegzés

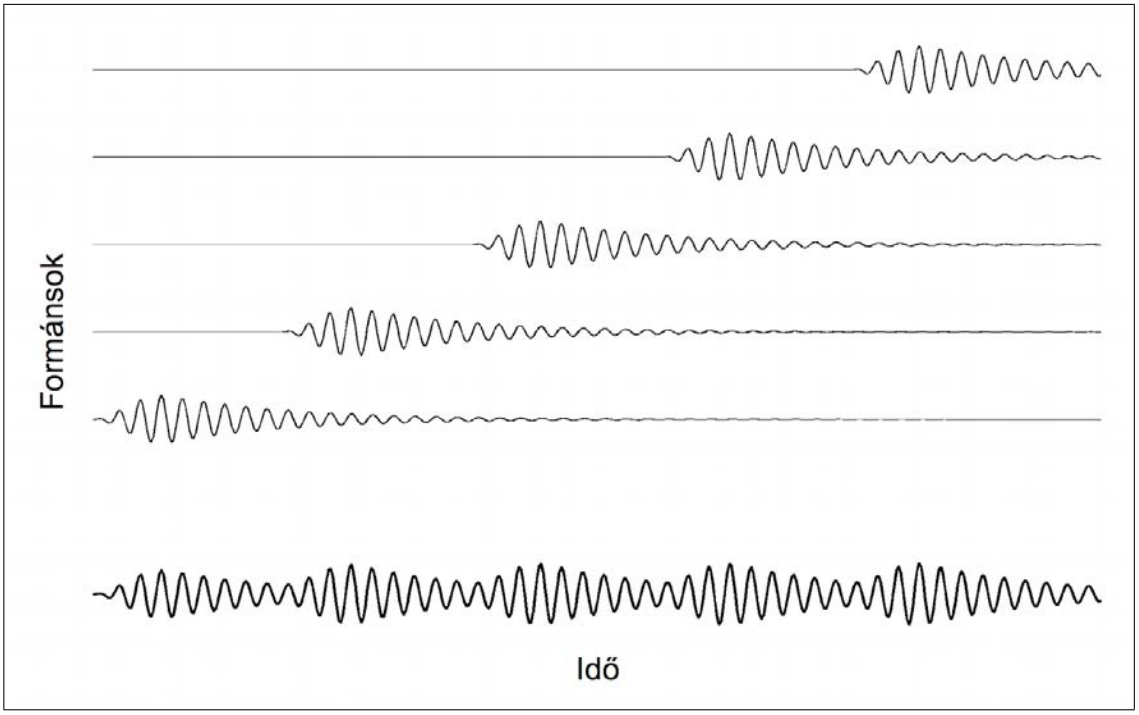
Hangszín. A módszer elsősorban kitartott hangszínek (magánhangzók, fúvós hangszerek) előállítására használható, habár a szintézis-paraméterek változtatásával interpolálni is lehetséges e hangszínek között. Zajos hangszínek szintézisére a módszer alkalmas, de nem túl hatékony.

Paraméterek. Az egyes formánsok elsődleges paraméterei zeneileg átláthatóak, a paraméterek folytonos változtatásával a hangszín folytonosan változik. Konkrét hangszínek imitálása komoly előzetes analízist, modellezést igényel. A másodlagos paraméterek, melyek a konkrét megvalósításban játszanak szerepet (pl. áteresztési arány vagy maximális szemcseszám), erősen befolyásolhatják az eredményt, habár zenei szerepük nehezebben fogható meg.

Műveletigény. Erősen implementációfüggő. Azonos erőforrás-felhasználás mellett általában az additív implementáció szól tisztábban, bár kifejezetten szűk sávok esetén a szubtraktív megoldás „olcsóbb”.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



7.10. ábra. A FOF-szintézis folyamata. Az alapfrekvencia által diktált pillanatokban egy-egy új hullámot generálunk (ezeket mutatják az egymás fölé helyezett formánshullámok), az így kapott hullámformákat pedig összeadjuk. Amint az ábrán is látható, általában az egyes alap-hullámformák később érnek véget, mint amikor a következő visszajátszás kezdődik, így „többszörös kiolvasással" kell élnünk.

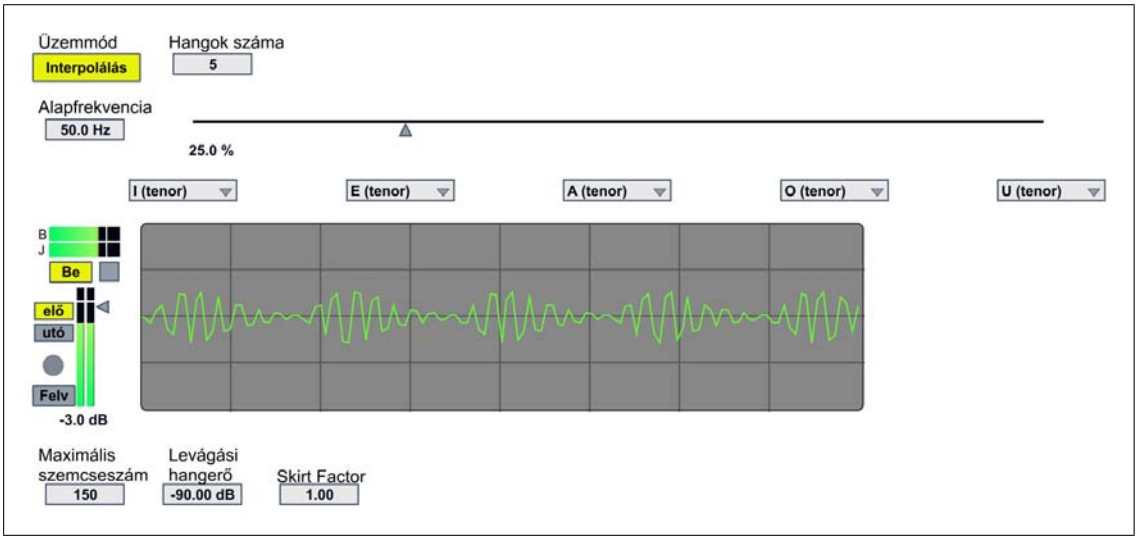
7.4. Feladatok

1. Hozz létre különböző formánsokat! Ehhez először nyisd meg a `HApp_07_03` `(OSX)` `(WIN)` példát és másold le legalább három különböző hang formánsait. Ezután nyisd meg a `HApp_07_02` `(OSX)` `(WIN)` példát és válaszd ki a középső modult. Az imént kimásolt adatok segítségével hozz létre olyan formánsokat, amelyek a legjobban közelítik a `HApp_07_03` `(OSX)` `(WIN)` példa hangjait! Vajon a kevesebb formáns, vagy a kizárólag egységesen vezérelhető sávszélesség okozza a nagyobb eltérést a két hang között? A kérdés eldöntéséhez olyan hangzókat válassz, ahol a legfontosabb komponensek sávszélessége közel azonos (pl. tenor „E").

2. Vizsgáljuk a `HApp_07_02` `(OSX)` `(WIN)` program elektronikus dorombját. Mi változik a hangszínben, ha a három csatorna közül kettőt kiiktatunk (pl. úgy, hogy -120 dB-re állítjuk)? Miért? Befolyásolja-e a hangszínt (és ha igen, akkor milyen módon) az egyes formánsok sávszélessége?

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



7.11. ábra. `HApp_07_04` `(OSX)` `(WIN)`: *FOF-szintézis interpolációval. Először beállítjuk a bal felső számdobozban a magánhangzók számát, majd az egyes legördülő listákból kiválasztunk egy-egy tetszőleges hangzót. A szintézis-paraméterek (alapfrekvencia, skirt factor, maximális szemcseszám, levágási hangerő) és a kimeneti hangerő beállítása után kiválasztjuk a kívánt üzemmódot (impulzusvezérelt, illetve folytonos). Impulzusüzemmódban az indítógomb megnyomásával, folytonos üzemmódban az interpolációs csúszka állításával hallhatjuk a szintézis eredményét. A jobb alsó részen levő oszcilloszkóp a szintetizált jel időbeli lefutását mutatja meg.*

Open Access 2026-11-02

8. fejezet

Modulációs technikák I: amplitúdó- és ringmoduláció

A jelmoduláció eredeti formájában a telekommunikáció területén játszott fontos szerepet, hiszen a rádiójelek átviteléhez fejlesztették ki.

Moduláció akkor történik, ha egy oszcillátor paraméterét egy másik jel időben változtatja, modulálja. A legáltalánosabb modulációs technikák az amplitúdó-moduláció (AM) és a frekvencia-moduláció (FM). Mindkettő jellemzője, hogy kevés kiindulási jel használatával jól kiszámítható, a kiindulási jeleknél összetettebb spektrumot képesek létrehozni.

A modulációban két jel vesz részt: a vivő- és a modulálóhullám. A vivőhullám modulálatlan (alapesetben: szinuszos) jel, amelyet a modulálójel változtat.

E fejezetben az amplitúdó-modulációt – valamint annak egy speciális esetét, a ringmodulációt[1] (RM) – mutatjuk be. A frekvencia-modulációval a következő fejezetben ismerkedünk meg.

8.1. Elmélet

8.1.1. A folyamat

Az AM és RM során a modulálójel a vivőjel amplitúdóját változtatja. Az amplitúdó-moduláció során a moduláló jelhez egy konstans amplitúdót adunk, az így kapott jellel pedig a vivőhullám amplitúdóját vezéreljük. A ringmoduláció az amplitúdó-modulációnak az a speciális esete, amikor a hozzáadott konstans jel értéke zérus. A két folyamatot a 8.1. ábra illusztrálja.

Tartott hang amplitúdójának lassú modulálásakor a zenében gyakran alkalmazott tremolóhoz hasonló, pulzáló hangosság keletkezik. Ha gyorsítjuk a modulálást, a pulzálás először sűrűsödni kezd, később összefüggő, új színezettel rendelkező folyamatos hanggá válik, majd egy-egy hangmagasság elkezd lefelé és felfelé glisszandálni.

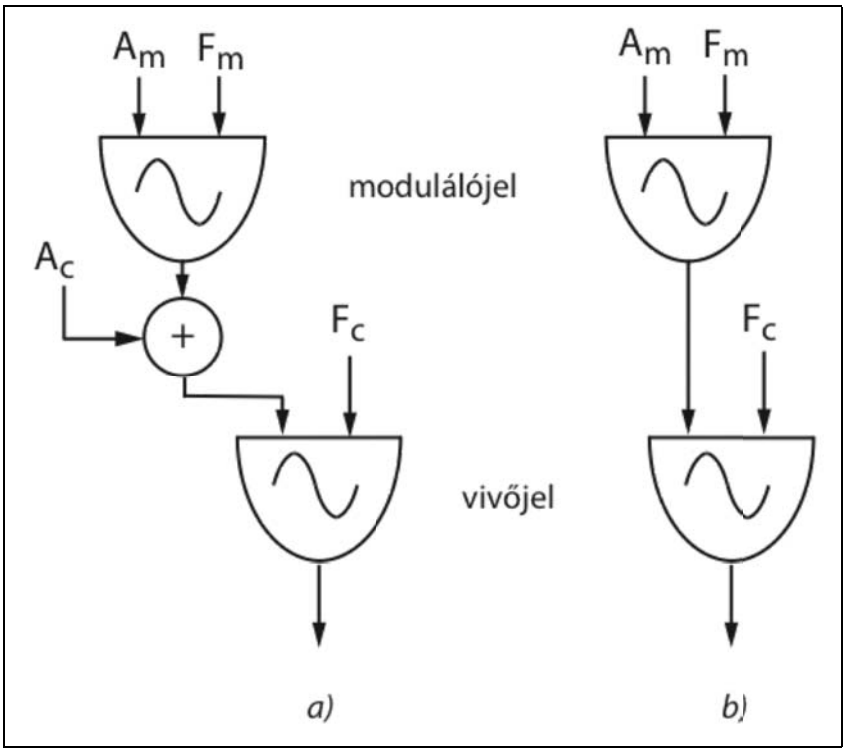
8.1.2. A spektrum és a hangszín

A két moduláció esetén hasonló spektrumok keletkeznek. Amennyiben egyszerű modulációt használunk, azaz a kiindulási jeleink szinuszhullámok, amplitúdó-moduláció esetén a

[1] Nevét az első analóg implementációk jellegzetes, gyűrűre emlékeztető kapcsolási rajzáról kapta.

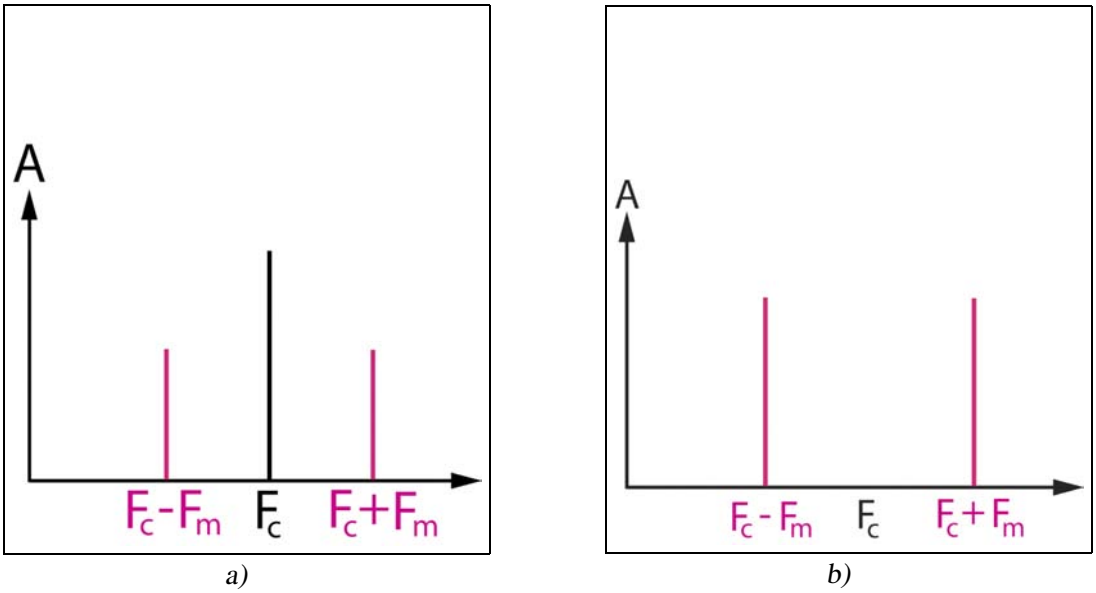
© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



8.1. ábra. Az amplitúdó- (a) és ringmoduláció (b) folyamatábrái.

vivőfrekvencia (f_c) mellett mindkét irányban egyenlő ($\pm f_m$) távolságra megjelenik egy-egy oldalfrekvencia. Ringmoduláció esetén f_c hiányzik a spektrumból, ilyenkor csak két helyen (az $f_c \pm f_m$ frekvenciákon) találhatóak összetevők. A két spektrumot a 8.2. ábra illusztrálja.



8.2. ábra. Az amplitúdó- (a) és ringmoduláció (b) spektruma.

f_c és f_m különbsége meghatározza a keletkező hangszín típusát. Az amplitúdó- és a ringmoduláció szinuszhullámok használatával nagyon egyszerű spektrumokat eredményez. Összetett hangzások elérése érdekében legalább a hordozó jelnek bonyolultabb hangszínnel

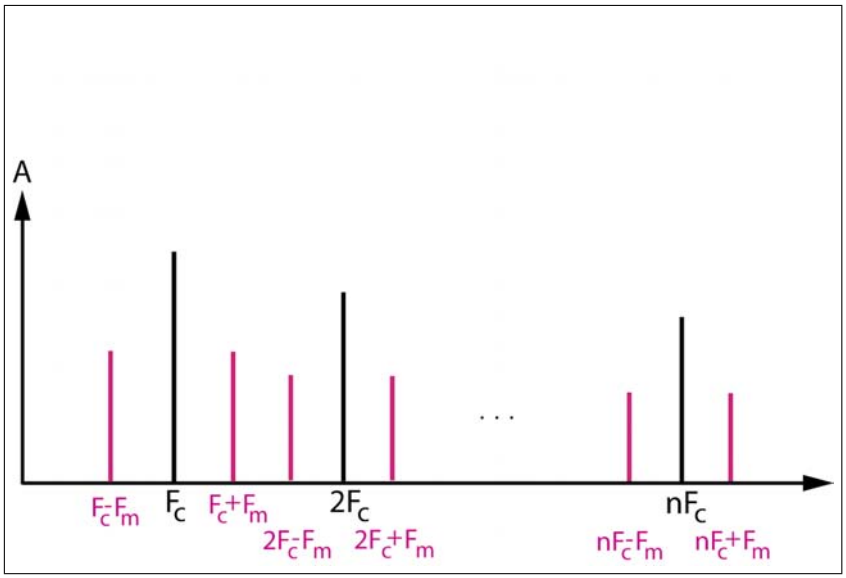
© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

kell rendelkeznie. Komplex hordozójel modulációja esetén a spektrum minden komponense mellett megjelennek az $f_i \pm f_m$ oldalfrekvenciák (itt f_i a spektrum i-edik komponensének frekvenciája). A 8.3. ábra harmonikus hang amplitúdó-modulációja esetén keletkező spektrumot illusztrálja. Az összetevők a vivőjel alapfrekvenciájának (f_c) felhangjai. Az oldalfrekvenciák megjelenése dúsítja a spektrumot; az $\frac{f_c}{f_m}$ aránytól függően különböző típusú hangzások keletkeznek.

Alapvetően négy jelleg alakulhat ki:

- Ha $0 < f_m \lessapprox 10$ Hz: lebegő, tremoláló hang.
- Ha $10 \lessapprox f_m \lessapprox 20$ Hz: érdes hang, melynek alaphangmagassága megegyezik a hordozóéval.
- Ha f_m és f_c valamilyen frekvencia egész számú többszörösei: színezett harmonikus hang.
- Ha a fentiek közül egyik sem teljesül: inharmonikus, több magasságérzetet keltő hang.



8.3. ábra. Amplitúdó-moduláció komplex vivőhullámmal.

A_c és A_m arányának folyamatos változtatásával interpolálni lehet az eredeti hang és a két technika alkalmazásával létrehozott hangzások között.

Mivel a ringmodulált hangzás csak a moduláció által létrehozott összetevőket tartalmazza (az eredeti hang spektrumát nem), ezért a ringmoduláció segítségével lényegesen messzebbre lehet távolodni az eredeti hangszíntől. Például, ha a moduláló szinusz frekvenciája is és a modulált hang alaphangja is ugyanannak a frekvenciának két nagyon alacsony felhangja (pl. a duplája és a triplája), a ringmoduláció transzponálja az eredeti hangot.

A különböző eshetőségeket a 8.1. táblázatban található hangpéldák illusztrálják.

Megjegyezzük, hogy amennyiben valamelyik létrejövő frekvencia negatív értéket venne fel, vagy (digitális implementációk esetén) meghaladná a mintavételi frekvencia felét, az adott frekvenciaértéket „tükrözni” kell. Mivel ez a jelenség a frekvencia-moduláció során nagyobb szerepet játszik, ezért a 9.1.3. alfejezetben ismertetjük.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

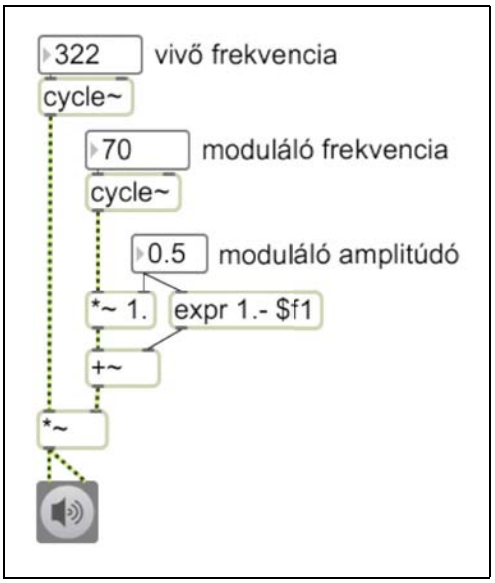
Open Access 2026-11-02

8.1. táblázat. AM és RM hangpéldák. Az első oszlop a vivőjel típusát határozza meg (szinuszoid/zongora). A második sorban levő sebességek a modulálójel frekvenciájára vonatkoznak.

	AM			RM		
	lassú	**közepes**	**gyors**	**lassú**	**közepes**	**gyors**
Szinuszoid	8.1	8.2	8.3	8.4	8.5	8.6
Zongora	8.7	8.8	8.9	8.10	8.11	8.12

8.2. Megvalósítás

Az AM és RM egyik előnyös tulajdonsága, hogy nagyon egyszerűen megvalósíthatóak. A technika legegyszerűbb implementációjának (szinusz modulál szinuszt, burkológörbék használata nélkül) CSound és Max programnyelveken készített implementációit a 8.1. forráskód és a 8.4. ábra mutatja be (a 8.13. hangpélda tartalmazza a CSound programból lekevert hangot).



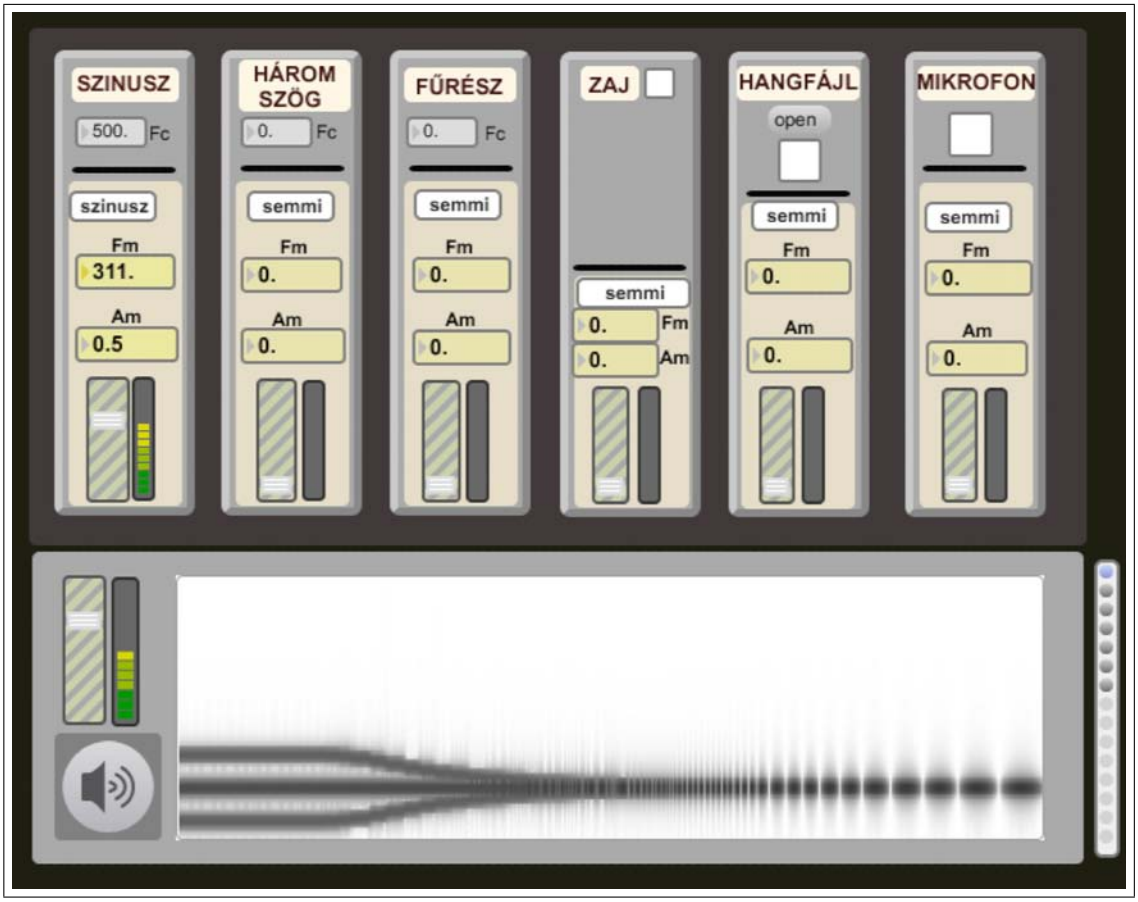
8.4. ábra. Amplitúdó-moduláció Max programnyelven (`HApp_08_01` `(OSX)` `(WIN)`).

A 8.5. ábrán bemutatott interaktív példa segítségével ki lehet próbálni, hogyan szólnak az egyszerű és összetett AM és RM különböző változatai. A vezérlőfelület 6 modulra oszlik, melyek 6 különböző hordozóhullámot kezelnek: szinusz-, háromszög- és fűrészfoghullámot, szűrővel kiegészített fehérzajt, beolvasható hangállományokat és mikrofon (vagy vonal) bejövő jelét. Minden modulban 3 modulálóhullám (szinusz-, háromszög- és fűrészfogjel) közül lehet választani. Állítható a periodikus hordozójelek frekvenciája (zaj, hangfálj és mikrofonjel esetében a bejövő jel összetevőinek frekvenciái válnak a vivőjellé), a moduláló frekvencia és a modulálójel amplitúdója.

A 8.5. ábrán az első modul (szinusz hordozót szinuszjel modulál) működése figyelhető meg, miközben a modulálófrekvenciát 311 és 1 Hz között változtattuk (lásd a 8.14. hangpéldát). A szonogramon követhető, ahogyan időben változik a moduláló jel frekvenciája: először egy „konstans” akkord szól, majd a két oldalfrekvencia emelkedését, illetve ereszke-

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



8.5. ábra. Amplitúdó-moduláció Max programnyelven.

dését követően egy érdes alapfrekvenciát hallunk. A moduláló-frekvencia további csökkenésével az érdes hanghatás tremolová válik, a hang lebeg.

Hangfájl beolvasásához az 5. modulban először be kell olvasni a kívánt hangfájlt az „open” gomb megnyomásával. A beolvasás után a zöld színű kapcsolóval lehet elindítani a lejátszást. A sárga kapcsolóra kattintva loopolni lehet a lejátszást.

8.3. Összegzés

Hangszín. A végeredmény a beérkező jeltől függ. Lehetséges tremolálóvá, illetve érdessé tenni ezt a jelet. Emellett megfelelő paraméterezéssel a spektrum is transzformálható.

Paraméterek. Kevés paraméterrel kell dolgozni, hiszen a jelek frekvenciái, amplitúdói és – AM esetén – az eredeti jel aránya a végeredményben elegendőek a hangzás meghatározásához. Ám mivel a létrejövő hangszín erősen függ a frekvenciák arányától (pontosabban, annak racionális voltától), ezért a módszer – amennyiben nem tremoloként használnánk – nem eléggé intuitív: csak hosszas előzetes számolással határozható meg, hogy adott paraméterek milyen hangszínváltozást okoznak.

Műveletigény. Az eljárás gyakorlatilag egyetlen szorzást igényel, ezért elhanyagolható az erőforrásigénye.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

8.4. Feladatok

1. Vegyél fel egy harmonikus hangot (pl. ének) egy mikrofon segítségével a HApp_08_01 (OSX) (WIN) jobb oldalán található panel használatával, majd...
   - ...adj hozzá tremolo-hatást!
   - ...tedd a hangot érdessé!
   - ...transzponáld harmonikusan és inharmonikusan!

   Ismételd meg a feladatot több (különböző jellegű) felvett hanggal!

2. Tölts be egy zongorahangot a HApp_08_01 (OSX) (WIN) hangfájl-lejátszó moduljába. Ebből kiindulva hozz létre legalább 3 különböző, preparált zongorára emlékeztető hangzást! Készítsd el továbbá a hang 3 különböző (harmonikus) transzpozícióját is!

8.5. Függelék

8.1. CSound-forráskód. hd_08_01_csound.csd

```
<CsoundSynthesizer>
<CsInstruments>

sr      = 44100
kr      = 4410
ksmps   = 10
nchnls  = 1

instr 1
        aamod       oscili      p6, p7, 1        ; modulálójel
        asig        oscili      p4+aamod, p5, 1  ; vivőjel
        out         asig
endin

</CsInstruments>
<CsScore>

        f1 0 2048 10 1
        i1 0 3 1500 300 1500 100
        e

</CsScore>
</CsoundSynthesizer>
```

Open Access 2026-11-02

9. fejezet

Modulációs technikák II: frekvencia-moduláció

Az előző fejezetben ismertettük az amplitúdó-modulációt, ahol a modulálójel a vivőjel amplitúdóját befolyásolja. E fejezet a frekvencia-modulációt (FM) mutatja be, mely során a vivőjel frekvenciája függ a modulátortól.

9.1. Elmélet

9.1.1. A folyamat

Az FM során egy egyszerű hullám (klasszikus esetben szinusz) frekvenciáját egy másik jel (klasszikus esetben szintén szinusz) modulálja.[1] Eredményként egy lényegesen összetettebb hang jön létre, mint a modulációban részt vevő hullámformák. Ahogyan a 9.1. ábrán látható, frekvencia-modulációkor a vivőhullám frekvenciájához hozzá kell adni egy szinuszfüggvény szerint változó jel értékeit.

Minél nagyobb a modulálójel amplitúdója, annál nagyobb a vivőjelben az eltérés a moduláció hatására létrejövő legnagyobb és a legkisebb frekvencia között. Ezt az eltérést *frekvencia-kitérésnek*, vagy *frekvencia-deviációnak* nevezzük (f_Δ). Ez jól hallható abban az esetben, ha kicsi a moduláló frekvencia, hiszen ekkor a frekvencia-modulációt vibratoként lehet érzékelni, és jól hallható a frekvencia-kitérés változása, amint az a 9.2. ábrán látható. A 9.1. hangpéldában egy 6 Hz frekvenciájú jel modulálja a vivőjelet.

A módszert a következő paraméterekkel határozzuk meg:

Vivőfrekvencia: a modulált oszcillátor frekvenciája (f_c).

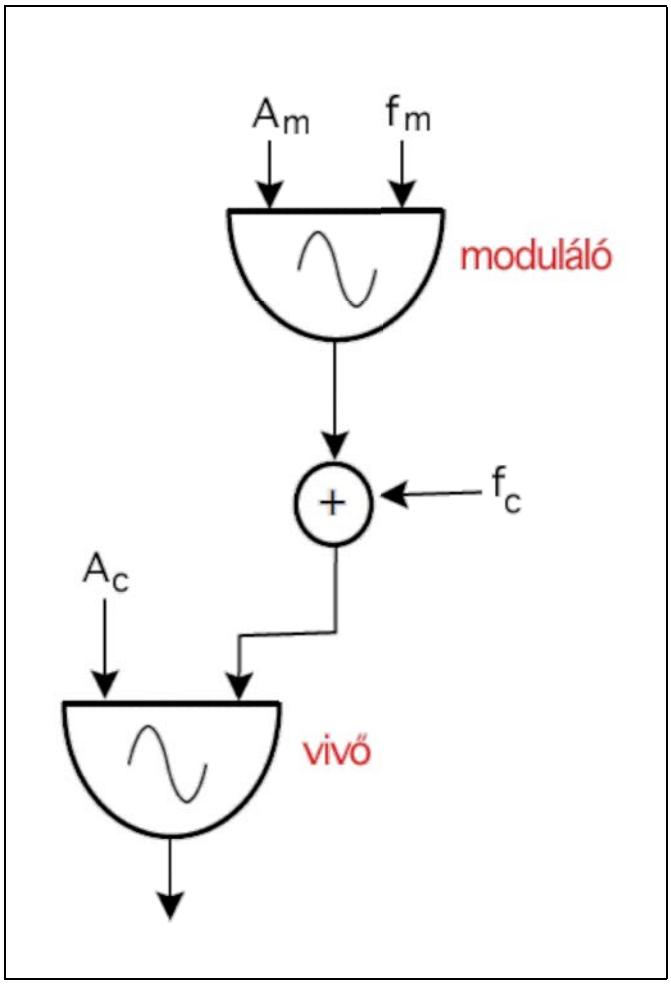
Moduláló frekvencia: a modulálójel frekvenciája (f_m).

Modulációs index: a frekvencia-deviáció (f_Δ) és a moduláló frekvencia aránya $\left(I = \frac{f_\Delta}{f_m}\right)$. A frekvencia-deviáció (vagy frekvencia-kitérés) a lehetséges legnagyobb eltérés az eredeti vivőfrekvencia és a végeredményként kapott jel pillanatnyi frekvenciája között. Értéke közvetlenül függ a modulálójel amplitúdójától.

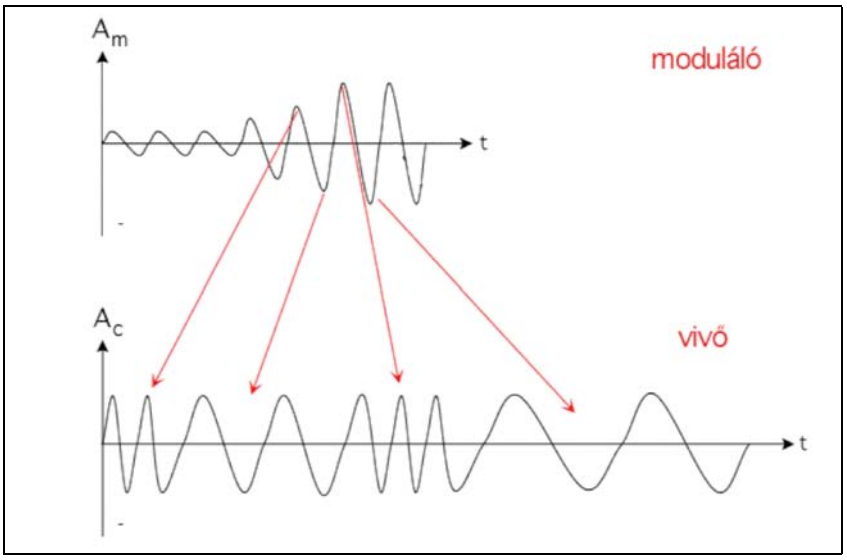
[1] A fejezet hátralevő részében feltételezzük, hogy mind a vivő- mind a modulálójel egyaránt szinuszos.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



9.1. ábra. A frekvencia-moduláció folyamatábrája.



9.2. ábra. Frekvencia-deviáció a modulálójel amplitúdójának függvényében.

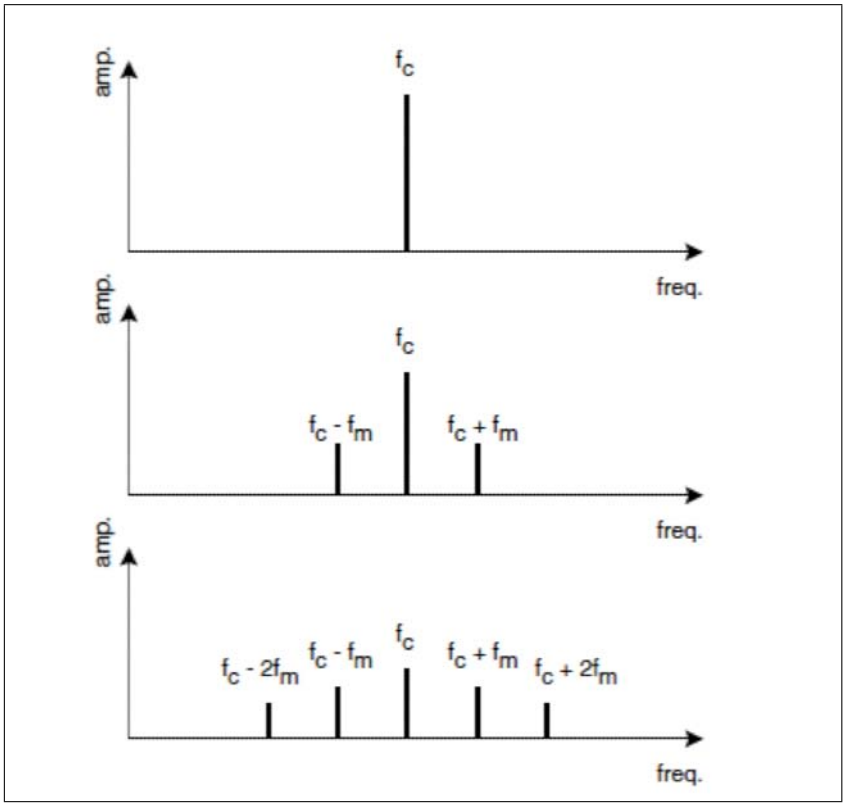
A modulációs index komoly szerepet játszik a frekvencia-modulált hangzás spektru-

Open Access 2026-11-02

mának kialakulásában. Látható, hogy minél nagyobb a modulálójel amplitúdója (és így a frekvencia-kitérés), annál nagyobb a modulációs index.

9.1.2. A spektrum és a hangszín

A frekvencia-moduláció során f_c körül végtelen sok oldalsáv jelenik meg, f_c-re szimmetrikusan. Az oldalsávok frekvenciái $f_c \pm k f_m$ alapján számolhatóak, ahol k a pozitív egész számokon fut végig. A kialakuló struktúrát a 9.3. ábra szemlélteti. Fontos kiemelni, hogy mivel végtelen sok pozitív egész létezik, ezért mindig végtelen sok oldalsáv jelenik meg.



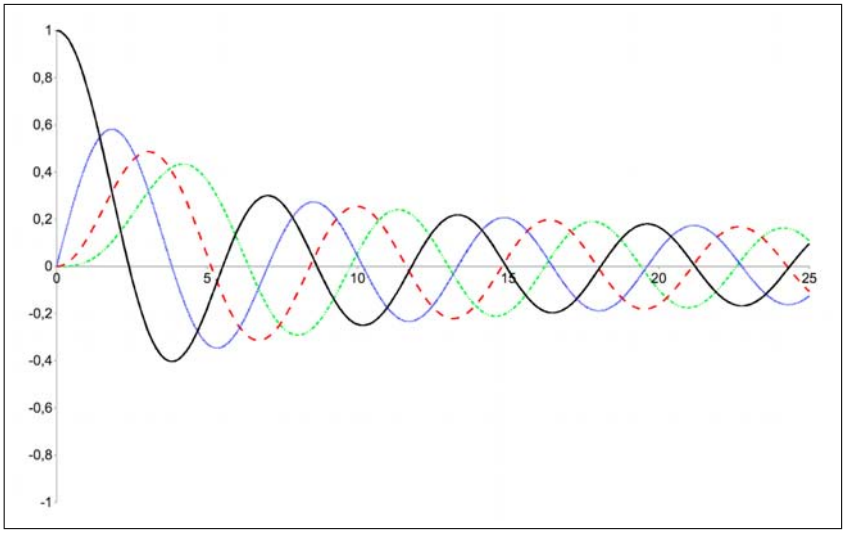
9.3. ábra. Oldalfrekvenciák megjelenése.

Az egyes oldalsávok amplitúdóját az oldalsáv sorszáma (k) és a modulációs index együttesen határozzák meg, az elsőfajú Bessel-függvényeken keresztül. A k-adik amplitúdó konkrét értéke $A_k = \pm J_k(I)$, ahol J_k a k-adik elsőfajú Bessel-függvény (az első néhány Bessel-függvényt a 9.4. ábra mutatja). Az amplitúdó előjelét az határozza meg, hogy az oldalsáv f_c felett $(+)$ vagy alatt (páratlan k esetén $-$, páros k esetén $+$) helyezkedik el. Az eredeti vivőfrekvencia végső amplitúdóját az $A_0 = J_0(I)$ összefüggésből származtatjuk.

Ahogy a 9.4. ábrán láthatjuk, a Bessel-függvények növekvő modulációs index mellett oszcillálva ugyan, de folyamatosan csökkennek. Ezért f_c-től távolodva az egyes oldalsávok amplitúdói csökkenő tendenciát követnek a Bessel-függvények oszcillációinak erejéig. Bizonyítható, hogy azok az oldasávok, amelyekre $k > I + 1$ teljesül, elhanyagolható módon

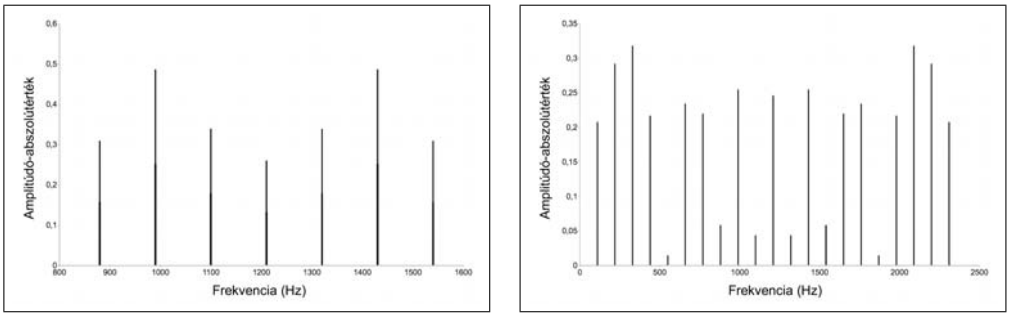
© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



9.4. ábra. A hordozójel és az első három oldalfrekvencia amplitúdójához tartozó Bessel-függvény. A fekete vonal a 0., a kék az 1., a piros a 2., a zöld a 3. elsőfajú Bessel-függvényt mutatja.

befolyásolják a hangszínt,[2] ezért a spektrum meghatározásakor ezektől eltekinthetünk. Ez azt is jelenti, hogy a modulációs index meghatározza a spektrum szélességét, amit a 9.5. ábrán illusztrálunk.



9.5. ábra. A modulációs index hatása a spektrum szélességére. Mindkét esetben $f_c = 1210$ Hz és $f_m = 110$ Hz (vagyis $H = \frac{1}{11}$, a hang harmonikus), azonban míg a bal oldali ábrán $I = 3$, a jobb oldalin $I = 10$. A két hangot a 9.2. (bal oldali spektrum) és a 9.3. (jobb oldali spektrum) hangpéldák tartalmazzák.

A módszer lehetőséget ad harmonikus és inharmonikus spektrum létrehozására egyaránt.

Harmonikus spektrum akkor keletkezik, ha a két frekvencia, f_m és f_c aránya racionális (az olyan számokat hívjuk racionálisnak, amelyek felírhatóak két egész szám hányadosaként). Ha az arányuk tovább nem egyszerűsíthető alakja[3] $\frac{p}{q}$, akkor a kapott hang f_0 alapfrekvenciáját az a szám határozza meg, amivel bővíteni kell a tovább nem egyszerűsíthető alakot

[2] A 9.4. ábrán látható, hogy k növelésével a függvény „periódusa” (a zéruskereszteződések helye) is növekszik. Mivel $k = 0$ esetén a függvények 0-ból indulnak, ezért – fix I mellett – k növelésével $J_k(I)$ egyre csökken.

[3] Ez a racionális számok legelterjedtebb felírási módja. A nem egyszerűsíthető alak lényege, hogy nem találhatunk olyan (egész) számpárt, amivel a törtet fel tudnánk írni úgy, hogy annak nevezője csökkenjen.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

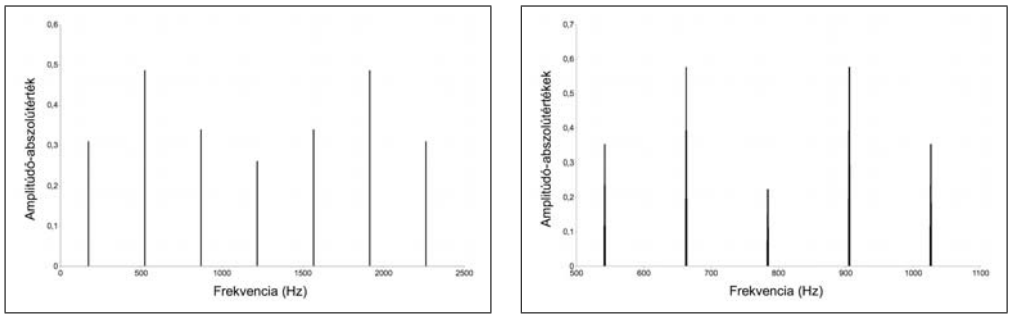
Open Access 2026-11-02

ahhoz, hogy az eredeti törtet kapjuk:

$$H = \frac{f_m}{f_c} = \frac{f_0 p}{f_0 q} \qquad \text{amiből} \qquad f_0 = \frac{f_c}{q} = \frac{f_m}{p}.$$

Az itt bevezetett H mennyiség a *harmonicitási arány*. A harmonikus hangzások alapesetei:

- Ha $q = 1$: a kapott hang egy olyan harmonikus hang, amelynek az alaphangja f_c.
- Ha $p = 1$: a kapott hang egy olyan harmonikus hang, amelynek az alaphangja f_m.
- Ha p páros szám, a spektrum csak a páratlan felhangokat tartalmazza.
- Ha $p = 3$, minden harmadik felhang hiányozni fog a spektrumból.
- Az egyre növekvő páratlan p értékek egyre hiányosabb, „szellősebb” spektrumot eredményeznek.



9.6. ábra. Harmonikus (bal) és inharmonikus (jobb) spektrumok. A bal oldali ábrán $H = \frac{f_m}{f_c} = \frac{348\ Hz}{1218\ Hz}$*, melynek tovább nem egyszerűsíthető alakja* $\frac{p}{q} = \frac{2}{7}$*; az alaphangra ebből* $f_0 = 174$ *Hz adódik, hiszen ekkor* $2f_0 = f_m$ *és* $7f_0 = f_c$ *is teljesül. A jobb oldali ábrán* $H = \frac{f_m}{f_c} = \frac{121\ Hz}{784\ Hz} \approx \frac{\sqrt{5}-1}{8}$*, ami irracionális, így a spektrumhoz nem tartozik egyértelmű alaphang. A két hangot a 9.4. (harmonikus) és a 9.5. (inharmonikus) hangpéldák tartalmazzák.*

Inharmonikus spektrum jön létre, ha a H harmonicitási arány irracionális, azaz nem fejezhető ki két egész szám hányadosaként.

Ha a hordozójel f_c alapfrekvenciája változik, a spektrum szerkezetének megőrzéséhez a modulálójel f_m frekvenciáját is arányosan meg kell változtatni (így a harmonicitási arány nem változik). Ha f_m-et állandó értéken tartjuk f_c változtatása mellett, a spektrum (és így a hangszín) változni fog a vivőfrekvencia függvényében.

A modulációs index a hangszín fényességére van nagy hatással, hiszen a lényeges oldalsávok száma a modulációs indextől függ. A növekvő számú oldalsáv következtében egyre magasabb frekvenciák jelennek meg a spektrumban, amitől a hangzás egyre fényesebbé, élesebbé válik. Mivel az oldalfrekvenciák a mély összetevők felé is bővülnek, ezért (főleg magas frekvenciájú vivőjel esetén) előfordulhat, hogy a modulációs index növekedésével járó elsődleges érzet nem a fényesedés, hanem a mély frekvenciák megjelenése lesz. A modulációs index időbeni változtatása egyszerű, hatékony eszköz időben változó spektrumok

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

létrehozására. A modulációs indexet szabályozó burkológörbe természetessé teheti a hangzásokat, és segít egyes hangzásjellegek kialakításában (pl. magas frekvenciák megjelenése az amplitúdó növekedésével a rézfúvós hangokban).

Az egyszerű frekvencia-moduláció lehetőségei nagymértékben bővíthetőek, ha további szinuszjelekkel vagy modulációkkal gazdagítjuk a szintézis folyamatát. Ilyenkor modulálhatunk egy hordozót több modulálójellel, több hordozót egy modulálójellel, összeadhatunk moduláló párokat (párhuzamos moduláció), a vivőjelet modulálójelet is modulálhatjuk (kaszkád moduláció), vagy akár visszacsatolásokat is iktathatunk a folyamatba. Az összetett modulációk segítségével lényegesen finomabb részletességgel lehet kidolgozni az elérni kívánt spektrumot.

A frekvencia-moduláció sikerét a 70-es, 80-as években az a tulajdonsága okozta, hogy kis számú kiindulási jelből nagyon sok, az egyes komponenseket változatos kombinációkban tartalmazó spektrumot lehet létrehozni, valamint, hogy a moduláló hullám időbeni változása a spektrum összetevőit közösen vezérelve természetes hatású, dinamikus hangzást okoz.

9.1.3. Álfrekvenciák keletkezése

E jelenség akkor lép fel, ha (akár AM, RM vagy FM során) bizonyos (számított) frekvenciák negatív értéket vennének fel, vagy (digitális rendszereknél) amikor egy frekvencia túllépné a mintavételi frekvencia felét.[4] Ilyenkor „tükrözni” kell a frekvenciát: negatív frekvencia esetén az eredeti frekvencia abszolút értékét kell vennünk, a mintavételi frekvencia felének túllépésekor pedig úgy kapjuk a helyes frekvencia-értéket, ha az eredeti frekvenciát kivonjuk a mintavételi frekvenciából. Összefoglalva, ha f a tükrözendő frekvencia, akkor

$$\begin{aligned} f<0: &\quad f = \quad f \\ f>\tfrac{1}{2}f_{\mathrm{sr}}: &\quad f = \quad f_{\mathrm{sr}}-f, \qquad \text{ahol } f_{\mathrm{sr}} \text{ a mintavételi frekvencia.} \end{aligned}$$

Fontos, hogy tükrözéskor a tükrözött komponens amplitúdója előjelet vált. Ha véletlenül egy tükrözött frekvencia egybeesik egy eredetileg is jelen levő frekvenciával, a két amplitúdót össze kell adni (az előjeleket figyelembe véve, természetesen). Amennyiben a tükrözést követően az adott komponens továbbra is kívül helyezkedik a megengedett tartományon, a kapott frekvenciát újra tükrözni kell mindaddig, amíg a végeredmény pozitív (ám a mintavételezési frekvencia felénél kisebb) nem lesz. Egy konkrét FM spektrum esetén a tükrözés módszerét a 9.7. ábra szemlélteti.

A nulltengely körüli tükröződésnek fontos szerepe van az FM spektrum végső formájának kialakulásában, különösen akkor, ha a tükrözést megelőző „eredeti” spektrum harmonikus volt. Hogy megértsük ennek okát, gondoljuk meg, hogy a negatív frekvenciák tükörképei az $nf_m-(f_c \bmod f_m)$ alakot[5] veszik fel (n egy nemnegatív egész szám). Emiatt, ha az eredeti spektrum harmonikus (vagyis a harmonicitási arány racionális), két eset lehetséges:

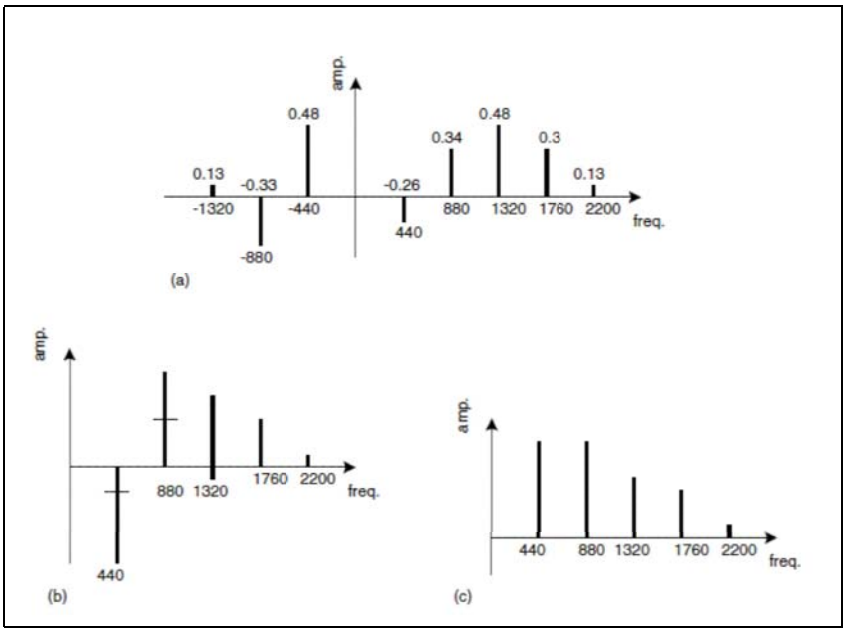
- Ha $\frac{2}{H}$ egész, akkor a tükrözött frekvenciák egybeesnek az eredeti spektrum pozitív frekvenciáival. Ilyenkor tehát nem jelennek meg új frekvenciák, ám az egyes komponensek amplitúdóinak előjelhelyes összeadása miatt a spektrum struktúrája – így a formánsok helye – átrendeződik. Ezt láthatjuk a 9.7. ábrán is.

[4] A mintavételi frekvencia felének e kitüntetett szerepe a Nyquist–Shannon tételből következik, lásd a 13.1.2. szakaszt.

[5] A mod a maradékképzés művelete: $f_c \bmod f_m$ azt a számot jelöli, amit maradékként kapunk, ha f_c-t elosztjuk f_m-mel.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



9.7. ábra. (a): Oldalfrekvenciák helyzete magas modulációs indexek esetén ($f_c = f_m = 440$). (b) és (c): oldalfrekvenciák pozitív tartományba hozása. Mivel $f_c = f_m$, ezért $H = 1$, így $\frac{2}{H} = 2$, ami egész. Emiatt a tükrözött és az eredeti frekvenciák átfedésbe kerülnek, és az előjelhelyes összeadás miatt az eredetileg 1320 Hz-nél levő formáns eltűnik; a spektrum új csúcsai 440 és 880 hertznél lesznek.

- Ha $\frac{2}{H}$ nem egész, akkor a tükrözött frekvenciák eltérnek az eredeti spektrum pozitív frekvenciáitól, így azok önálló komponensekként jelennek meg. Az, hogy a kapott hangszín harmonikus vagy inharmonikus, elsősorban f_m és $(f_c \bmod f_m)$ viszonyától függ. Szigorú értelemben véve, ha H racionális, akkor mind az eredeti, mind a tükrözés után kapott spektrum harmonikus lesz, ám előfordulhat, hogy a tükrözött spektrum túlságosan komplex struktúrája miatt a végeredmény inharmonikus benyomást kelt.

A két fenti esetet a 9.8. ábra illusztrálja.

9.2. Megvalósítás

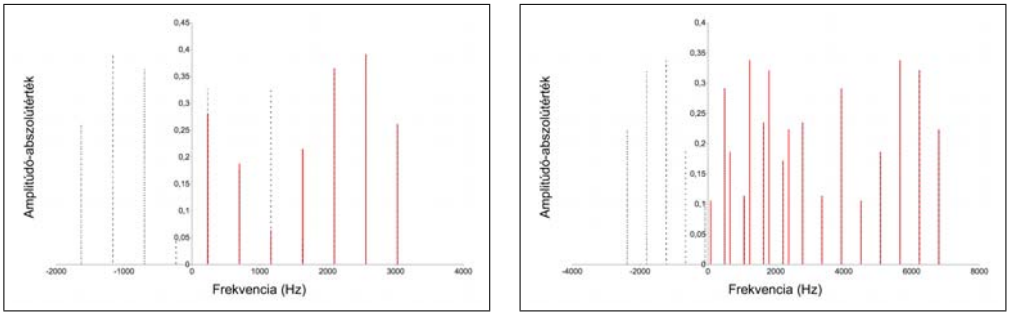
A frekvencia-modulációt a `HApp_09_01` `(OSX)` `(WIN)` programon keresztül ismertetjük, melyet a 9.9. ábra mutat. A példa két, egymástól teljesen független FM-blokkot, valamint egy spektrogramot tartalmaz.

A bal oldali modul egy egyszerű FM-szintetizátor, ahol önállóan állítható a hordozó- és moduláló-frekvencia, a modulációs index, valamint a szintetizált hang teljes időtartama és amplitúdója. Utóbbi egy burkológörbe segítségével modulálható.

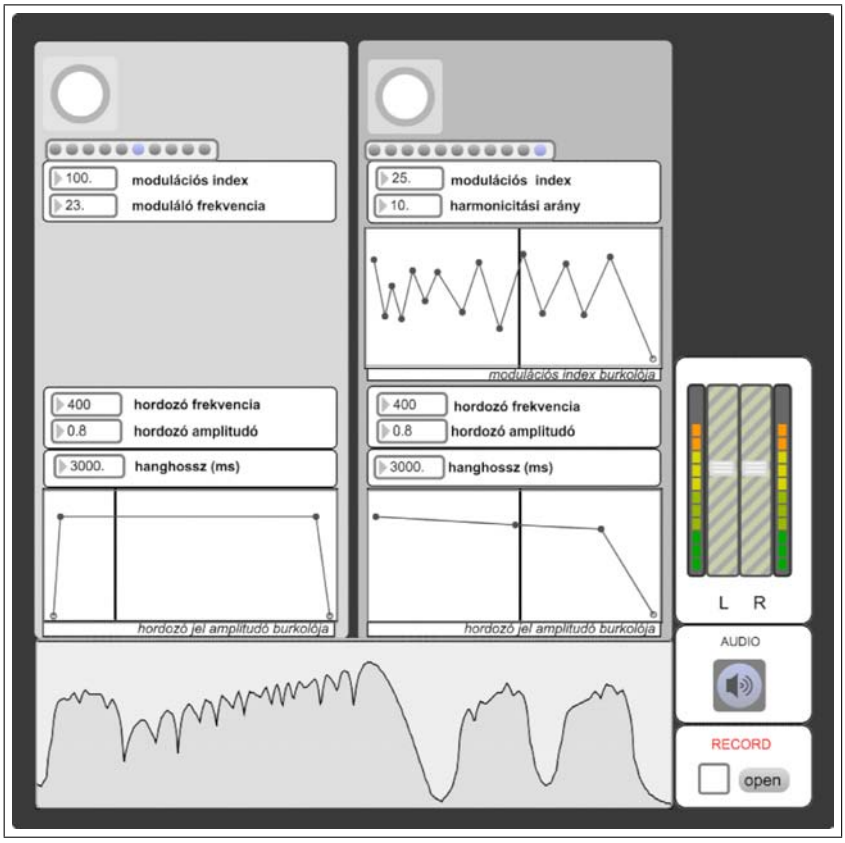
Az elmentett hangok vivőfrekvenciája egységesen 400 Hz, az amplitúdó-burkolók egyszerű trapézok. Az egyes hangok közötti eltérés így a moduláló-frekvenciában, illetve a modulációs indexben keresendő. Az első pár példán a különböző tremolokat figyelhetjük meg; látható, hogy a modulációs index növelésével a tremolo ambitusa, míg a moduláló-frekvencia emelésével a „sebessége” növelhető. A (balról számított) 6. hang (itt $f_m = 23$ Hz) már nem fogható fel tremolonak, ám még nem alakul ki stabil hangszín. A további beállítá-

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



9.8. ábra. Tükröződő spektrumok. A tükrözés előtti állapotot a pontozott fekete spektrum, a tükrözés utáni végeredményt a piros spektrum mutatja. A bal oldali ábrán $H = \frac{f_m}{f_c} = \frac{466\ Hz}{699\ Hz} = \frac{2}{3}$ *(ebből az alaphang* $f_0 = 233\ Hz$*) és* $I = 5$*. Mivel* $\frac{2}{H} = 3$*, várakozásainknak megfelelően a tükrözött frekvenciák az eredetiekkel átfedésbe kerülnek. A jobb oldali ábrán* $H = \frac{f_m}{f_c} = \frac{574\ Hz}{2214\ Hz} = \frac{7}{27}$ *(ebből az alaphang* $f_0 = 82\ Hz$*) és* $I = 8$*. Mivel* $\frac{2}{H} = \frac{27}{14}$*, a tükrözött frekvenciák új komponenseket illesztenek a spektrumba. Mivel* $f_c \bmod f_m = 492\ Hz$ *(ami megegyezik* $f_m - f_0$*-lal), a kapott hang, habár – matematikai értelemben – harmonikus, nagyon szélesen elhelyezkedő frekvencia-párokat tartalmaz, így nem alakul ki igazi hangmagasság-érzet. A két hangot a 9.6. (bal) és a 9.7. (jobb) hangpéldák tartalmazzák.*



9.9. ábra. Egy kétcsatornás FM szintetizátor. A jobb oldali ág lehetővé teszi a modulációs index időbeli változtatását (egy burkológörbe segítségével), a modulációs frekvencia helyett pedig a harmonicitási arányt használja.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

sok már eltérő hangszínű hangokat produkálnak, az előző szakaszban leírt elméleti megfontolások szerint.

A jobb oldali modulban a moduláló-frekvencia helyett a harmonicitási arány adható meg (ebből a moduláló-frekvencia az $f_m = Hf_c$ összefüggés alapján számolható), a modulációs indexet pedig egy burkológörbe modulálja. E példán keresztül könnyű igazolni a fejezet harmonicitási aránnyal és modulációs indexszel kapcsolatos megállapításait.

Az elmentett hangok elsősorban a modulációs index szerepét illusztrálják. Amint említettük, minél nagyobb a modulációs index, annál több komponens jelenik meg a spektrumban. A (balról) 7. hang esetén például a hang perkusszív karaktere nemcsak az amplitúdóburkológörbe jellegéből adódik, hanem a magas frekvenciájú komponensek gyors eltűnéséből is, ami a modulációs index csökkenésének köszönhető.

A 9.1. program egy egyszerű FM szintetizátort valósít meg CSound nyelven (lásd a 9.8. hangpéldát). Az első három hang ugyanazt a (konstans) vivőfrekvenciát modulálja eltérő (szintén konstans) modulátor-frekvenciákkal. A további hangok a modulációs index változtatásának a hatásait mutatják be.

9.3. Összegzés

Hangszín. Habár alacsony modulációs frekvenciák mellett vibrato előállítására alkalmas, az igazi előnye a módszernek, hogy a két nagy osztályon – a harmonikus és az inharmonikus spektrumokon – belül számos hangzástípust képes létrehozni. Habár nem képes annyiféle spektrumot létrehozni, mint az additív módszer, az alacsony paraméter- és számításigénye miatt nagyon elterjedt.

Paraméterek. Az oldalsávok lehetséges helyeit f_c és f_m, míg számukat a modulációs index határozza meg; mivel a spektrum ezen paraméterekből egyértelműen meghatározható, elmondhatjuk, hogy az FM nagyon kevés paraméter használatával képes igen sokféle és változatos hangszínt létrehozni. Sajnos, amint láttuk, a spektrum pontos meghatározása nagyon sok számolást igényel (jóval többet, mint szinte bármelyik szintézis- vagy modulációs technika esetén), ami jelentősen nehezíti a használatát. A legtöbb hangzást csak próbálgatással, kísérletezéssel lehet létrehozni. A paraméterek valós idejű változtatásával a hangszín jelentősen változik, ezért habár lehetséges interpolálni az egyes spektrumok között, a köztes lépésekben kialakuló hangszín nehezen jósolható meg előre.

Műveletigény. Habár nem olyan elenyésző, mint az amplitúdó-modulációé, erőforrásigénye e módszernek is minimális.

9.4. Feladatok

1. Számold ki és rajzold le, hogy milyen spektrum keletkezik, ha

   - $f_c = 100$ Hz, $f_m = 200$ Hz, $I = 5$
   - $f_c = 5$ kHz, $f_m = 200$ Hz, $I = 4$
   - $f_c = 200$ Hz, $f_m = 284$ Hz, $I = 6$

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

- $H = \frac{3}{5}$
- $H = \overline{3}$

Határozd meg, hogy melyik esetekben keletkezik harmonikus spektrum, és mi annak az alapfrekvenciája! Hozd létre a hangokat CSound-ban és/vagy Max-ben!

2. Határozz meg három különböző (f_c, f_m) párt, amelyek azonos alaphangú harmonikus hangokat hoznak létre!

3. Válassz ki három előre beállított hangot a `HApp_09_01` (OSX) (WIN) jobb oldali paneljéről! A harmonicitási index ismeretében határozd meg a spektrumok lehetséges frekvencia-komponenseit! Fogalmazd meg, hogy a modulációs index időbeli változása hogyan befolyásolja a spektrumokat az egyes konkrét esetekben!

9.5. Függelék

9.1. CSound-forráskód. `hd_09_01_csound.csd`

```
<CsoundSynthesizer>
<CsInstruments>

sr      = 44100
kr      = 4410
ksmps   = 10
nchnls  = 1

instr 1
        icarrierf  =          p5
        imodulatorf =         p6
        kcarenv    oscil1     0, 1, p3, p8
        kmodenv    oscil1     0, 1, p3, p9
        amodosc    oscili     p7*kmodenv, imodulatorf, 1
        acarrier   oscili     p4*kcarenv, icarrierf+amodosc, 1
        out        acarrier
endin

</CsInstruments>
<CsScore>


        f1 0 4096 10 1
        f2 0 4096  5 0.001  200 1   3696 1      200 0.001
        f3 0 4096  5 0.001    5 1   4095 0.001
        f4 0 4096  7 0     2096 1   2000 0
        f5 0 4096  7 1     4096 1
        f6 0 4096  7 2      200 0.5 1896 1     2000 0
        f7 0 4096  5 0.001   50 1   3801 1      245 0.001

        i1  0 2   10000  200    3  500 2 5
        i1  2 2   10000  200   20  500 2 5
        i1  4 2   10000  200  200  500 2 5
        i1  6 3   10000  200  200  800 2 4
```

Open Access 2026-11-02


```
        i1   9 6   10000   400   400 1500 7 6
        i1  15 0.1 10000   200   284 1500 3 3
        i1   +   .      .     >     >    > . .
        i1   +   .      .     >     >    > . .
        i1   +   .      .     >     >    > . .
        i1   +   .      .     >     >    > . .
        i1   +   .      .     >     >    > . .
        i1   +   .      .     >     >    > . .
        i1   +   .      . 1000  1300 2300 . .
        i1   +   .      .     >     >    > . .
        i1   +   .      .     >     >  200 . .
        i1   +   .      .     >     >    > . .
        i1   +   .      .     >     >    > . .
        i1   +   .      .     >     > 8000 . .
        i1   +   .      .   100   100  200 . .
        e

</CsScore>
</CsoundSynthesizer>
```

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

10. fejezet

Spektrumanalízis

Ahogy az 1. fejezetben említettük, a spektrumot úgy kapjuk, hogy a vizsgált hullámot szinuszos rezgések összegére bontjuk. Szót ejtettünk arról is, hogy a hangszín dimenziói (pontosabban, egy jól körülhatárolható részük) az így nyert spektrumból származtathatóak. Ebben a fejezetben a szinuszos felbontás részleteit vizsgáljuk meg.

10.1. Elmélet

Egy jel frekvencia-felbontását matematikailag a *Fourier-transzformáció* végzi el. A transzformáció egy időben végtelen kiterjedésű jel frekvencia-összetevőit határozza meg. A végeredmény egy – szintén végtelen kiterjedésű – függvény, ami minden egyes frekvenciaértékhez hozzárendeli az adott frekvencia *amplitúdóját* és *fázisát*. Utóbbi azt mutatja meg, hogy az adott frekvenciájú szinuszhullámot mennyire kell „eltolni" az időtengelyen – ha ezt az eltolást nem vennénk figyelembe, képtelenség lenne egy olyan összetett jelet felbontani, ahol az egyes frekvenciákat nem ugyanakkor „kapcsoltuk be". A transzformáció fontos tulajdonsága az *egyértelműség*: minden jelhez egyértelmű, egyedi spektrum tartozik. Ennek legfontosabb következménye, hogy a spektrum alapján egyértelműen rekonstruálható az eredeti jel. Ennek gyakorlati hasznát a következő szakaszban fogjuk látni.

A „végtelen kiterjedésnek" az a következménye, hogy – habár az egyes frekvenciákat pontosan meg tudjuk határozni – az egyes összetevők időbeliségéről semmit sem tudunk meg. Ez azt jelenti, hogy például egy két frekvenciát tartalmazó spektrum esetén, habár a két frekvenciát pontosan ismerni fogjuk, lehetetlen megmondanunk, hogy vajon e két frekvencia egyidejűleg szólalt meg, vagy egymás után. Ráadásul a gyakorlatban nem létezik „végtelenül hosszú idejű" hang. Ahogy az 1. fejezetben utaltunk rá, e problémákat úgy orvosoljuk, hogy a beérkező hangot rövid szakaszokra bontjuk, és az így kapott szakaszokat egy-egy végtelen hosszú jel részletének tekintjük. A kivágást úgy végezzük, hogy egy *ablakot* helyezünk a jelre, és eltekintünk a jelnek az „ablakon" kívül eső részétől. A módszer neve *ablakolás*.

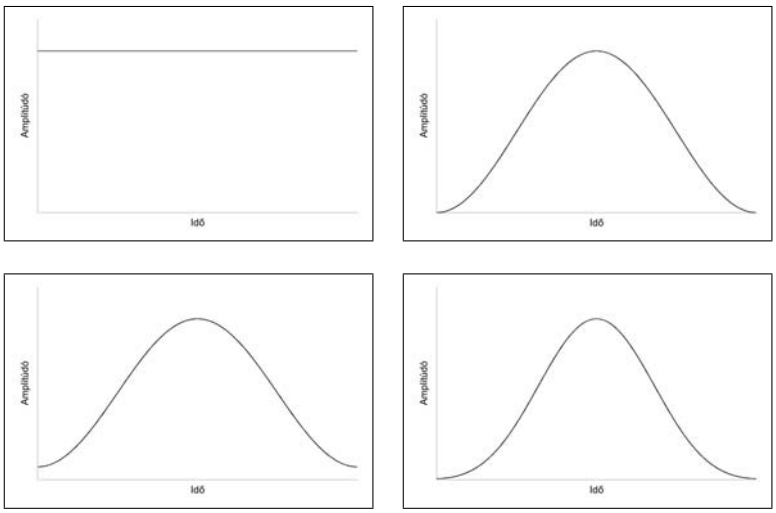
Sajnos azzal, hogy kivágunk egy szakaszt a „végtelen jelből", a frekvencia-elrendezést óhatatlanul torzítjuk.[1] A torzítás hatására olyan frekvenciák is megjelennek a spektrumban, amik az eredeti jelben nincsenek benne. Ennek kiküszöbölésére különböző *ablakfüggvényeket* alkalmazhatunk: az ablakolt jel amplitúdóját egy burkológörbével moduláljuk (néhány ismertebb ablakfüggvényt a 10.1. ábrán ismertetünk). Az ablakfüggvények alkalmazásának

[1] Ha nem így lenne, akkor a végtelen jel spektruma megegyezne a szakasz spektrumával, ami ellentmondana a transzformáció egyértelműségének.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

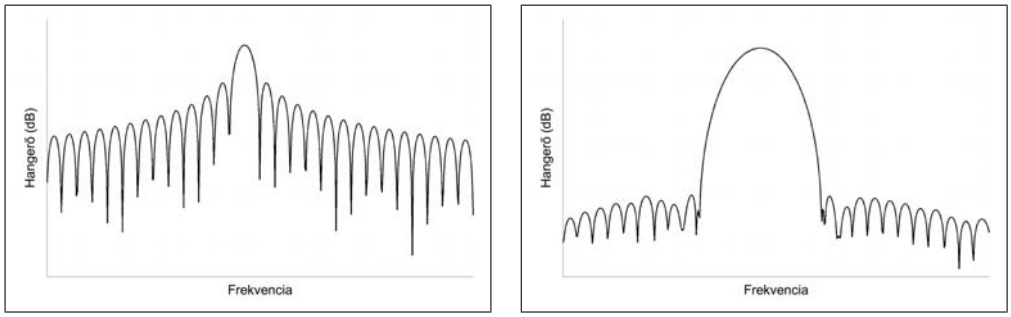
Open Access 2026-11-02

„negatív mellékhatása”, hogy az eredeti frekvenciákat többé-kevésbé elmossák: egy olyan jel spektruma, ami eredetileg egyetlen frekvenciát tartalmazott, egy kiterjedt sáv lesz az eredeti frekvencia körül. Így különböző ablakfüggvényeket szokás alkalmazni. Ezek egy része kevésbé mossa szét a spektrumot (ekkor azonban sok torzított frekvenciát is kapunk), egy másik részük a torzított frekvenciák számát minimalizálja (az eredeti frekvenciák nagyfokú „szétfolyatása” mellett). Amint láthatjuk, nincs királyi út.



10.1. ábra. Néhány a legismertebb ablakfüggvények közül. Felső sor: négyszögfüggvény *(bal),* Hann- *vagy* von Hann-függvény[2] *(jobb); alsó sor:* Hamming-függvény *(bal),* Blackman-függvény *(jobb).*

Az ablakfüggvények által okozott torzításokat a 10.2. ábra illusztrálja.



10.2. ábra. Szinuszhullám spektruma négyszögletes (bal) és Blackman–Harris (jobb) ablakfüggvények használatával (a „valódi” spektrum egyetlen frekvenciát tartalmazna, ahogy az az 1.1. ábrán látható). Míg a négyszögletes ablaknak a frekvenciafelbontása, addig a Blackman–Harrisnek a zajszintje kíváló.

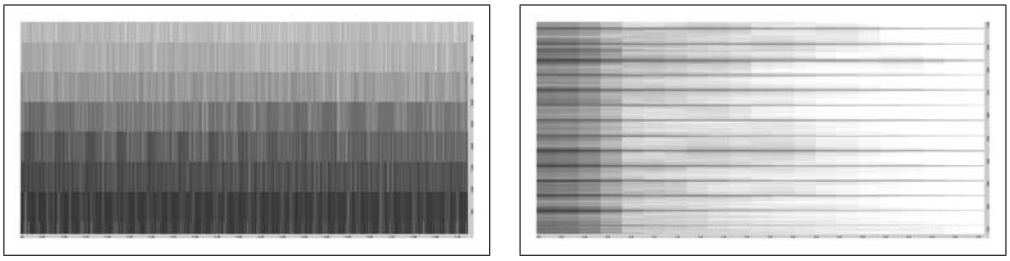
Maga a „szakaszokra bontás” is torzítja a végeredményt. Ennek az oka az idő és a frekvencia korábban már leírt „egylényegűségében” keresendő: mivel adott frekvenciához adott periódusidő tartozik, a kivágott szakaszból készített spektrum értelemszerűen nem tartalmazhat olyan frekvenciákat, amiknek periódusideje hosszabb, mint a kivágott szakasz időtartama. A *kvantálás* további problémát okoz: habár matematikailag a Fourier-transzformáció

[2] *Julius von Hann* osztrák meteorológus után. A közhiedelemmel ellentétben *nincs* Hanning-függvény!

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

a frekvenciákat folytonosan (végtelenül finom felbontásban) jeleníti meg, a gyakorlatban nem léteznek „végtelenül finom” mérőeszközök, ezért a transzformáció gyakorlati megvalósítása során *frekvenciasávokat* kapunk. Más szóval, csak azt tudjuk meg, hogy egy adott frekvenciasávban előforduló összes frekvenciának *együttesen* mekkora az amplitúdója; arról nem kapunk információt, ami az egyes sávokon belül történik. E két tény (a kivágás időtartamából adódó frekvenciakorlát, illetve a kvantálás) következménye, hogy a gyakorlati alkalmazásokban a spektrumban csak egyetlen alapfrekvencia, illetve annak felhangjai fognak szerepelni. Emiatt a felbontás nem egyformán érzékeny minden oktávban: gondoljuk meg, hogy pl. a 20 Hz-es felbontás (50 ms-os ablak), habár a magas tartományban pár centnek felel meg, az egyvonalas „A” körül már majdnem egy félhangot jelent. Ezt illusztrálja a 10.3. ábra.



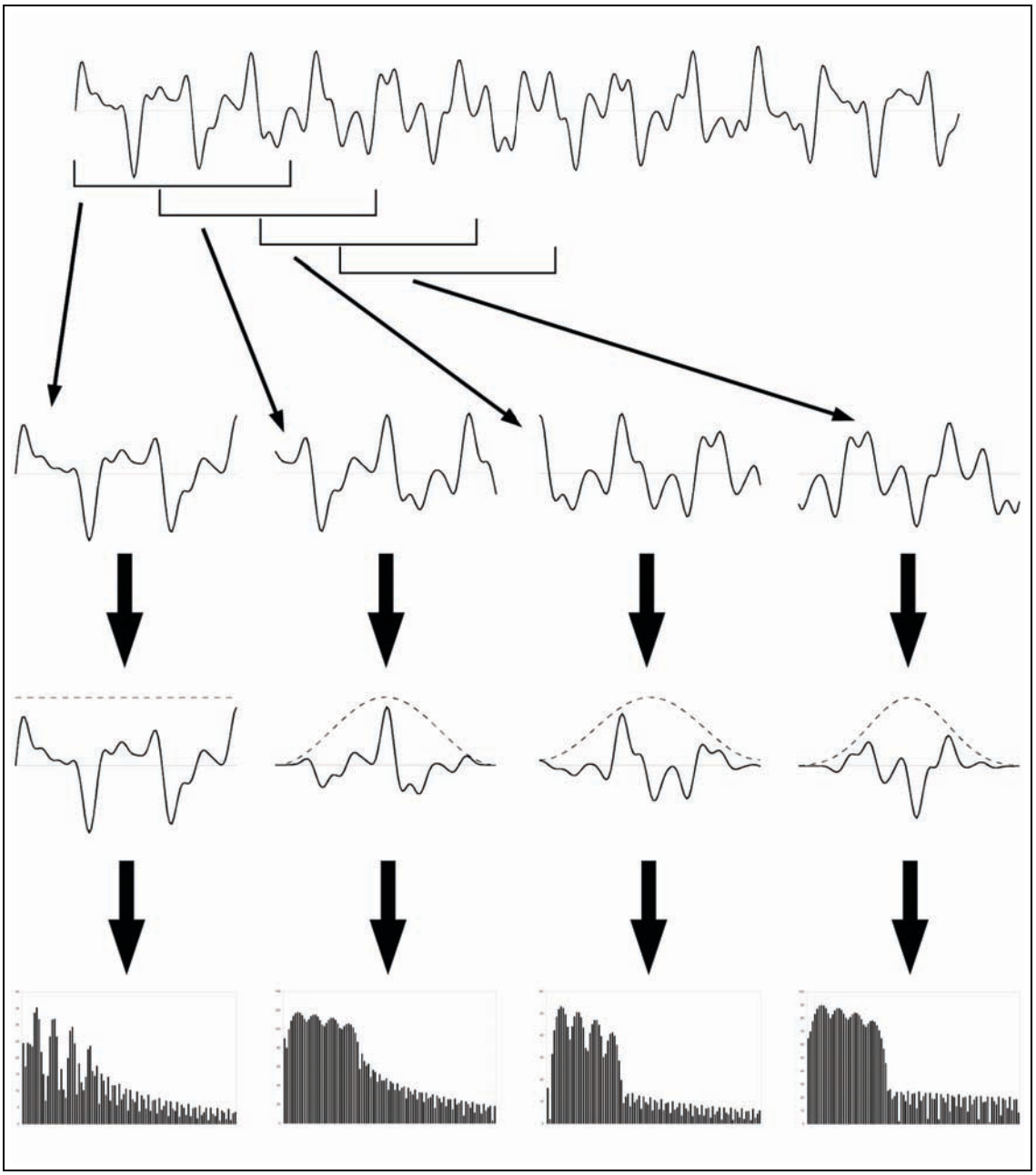
10.3. ábra. Az eltérő ablakméretek okozta torzítás vizsgálata. Mindkét szonogram a 10.1. hangminta alapján készült. A bal oldali szonogramon az ablakméret 1,45 ms (689 Hz-es felbontás), a jobb oldalin 1 486 ms (0,67 Hz-es felbontás). Látható, hogy míg a bal oldali szonogram időfelbontása kiváló, a frekvencia-felbontás annyira gyenge, hogy lehetetlen megmondani, milyen frekvenciák alkotják a spektrumot. A jobb oldali ábrával fordított a helyzet: itt a frekvencia-felbontás egyértelmű, ám az időfelbontás annyira rossz, hogy képtelenség megmondani, melyik frekvencia mikor szól. A szonogramok által megjelenített tartományok: 0–5 kHz és 1,5–1,7 s (bal); 100–1 000 Hz és 0–3,8 s (jobb).

A Fourier-analízis teljes folyamatát a 10.4. ábra ismerteti. Az ábra tetején látható az eredeti jel, melyet négy tiszta szinuszos hullámból kevertünk össze. A szakaszokra bontást az eredeti jel alatt elhelyezett kapcsok jelzik – a pontosabb felbontás érdekében ezek a szakaszok általában egymásba nyúlnak, így az analízis időfelbontását nem az egyes szakaszok időtartama, hanem az egymást követő kivágások közötti távolság határozza meg. A kivágott jelcsonkokat fentről a második sorban láthatjuk. Az ez alatti sor az ablakfüggvény alkalmazását mutatja. Habár egy konkrét analízis során végig egyetlen ablakfüggvényt használunk, az illusztráció kedvéért itt mindegyik jelszakaszt más-más ablakfüggvénnyel moduláltuk. Ezek balról jobbra: négyszögletes, Hann, Hamming és Blackman (a szaggatott vonalak jelzik az egyes ablakfüggvényeket, a folyamatos vonalak a modulált jelalakokat). Látható, hogy a legtöbb ablakfüggvény a jelet annak két szélén torzítja legjobban. A legalsó sorban a kapott ablakolt spektrumokat látjuk. Habár mindegyik spektrumon látható az eredeti négy szinusz, a végeredmények igen eltérnek. A négyszögletes ablaknál a csúcsok élesek, de sok „idegen” csúcs is megjelenik, ráadásul a jel/zaj arány nagyon rossz. A Hann ablak jól kihozza a négy csúcsot, de nagyon szétteríti őket, és a torzulások is széles sávon mozognak. A Hamming ablak az előző kettő közötti kompromisszumot kínál, míg a Blackman (habár a csúcsokat még a Hann-esetnél is jobban szétfolyatja) szinte alig ad zajt a spektrumhoz. Megjegyezzük, hogy a „helyes” spektrum négy vonal lenne a négy frekvencia helyén; láthatjuk tehát, hogy

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

mekkora mértékű torzítás keletkezik az ablakolás hatására.



10.4. ábra. A Fourier-analízis lépései.

Két gyakorlati kérdésre érdemes odafigyelni. Az egyik, hogy amennyiben az ablakolt jelet „meghosszabbítjuk” (csendet illesztünk az ablakolt jel után), a végeredmény „simábbá” tehető. A módszer, melynek neve *felülmintavételezés*, habár nem képes megnövelni a frekvencia-felbontást, megkönnyíti a spektrum további analízisét (pl. formánsok meghatározását). A másik, hogy egy speciális algoritmus nagyon gyorsan képes elvégezni a Fourier-transzformációt, ha a transzformálni kívánt adatsorban levő minták száma megegyezik 2 valamelyik pozitív egész hatványával (2, 4, 8, 16, 32 stb.). Az algoritmus neve *gyors Fourier-transzformáció* (*Fast Fourier-Transform, FFT*). A legtöbb program ezért (amennyiben a felhasználó által beállított ablakméret nem maga is egy 2-hatvány) felülmintavételezés-

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

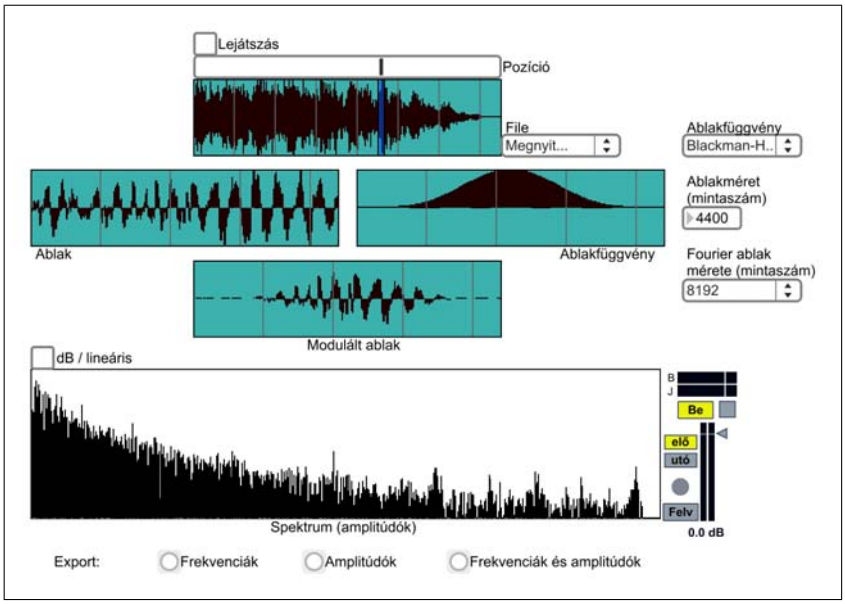
Open Access 2026-11-02

sel biztosítja, hogy magát a transzformációt mindig 2-hatvány hosszúságú ablakokon kelljen végezni.

Megjegyezzük, hogy nem a Fourier-transzformáció az egyetlen módszer egy jel felbontására. A legnépszerűbb alternatív eljárás, a *wavelet-transzformáció*, olyan apró hullámszemcsék összegére bontja a jelet, amilyeneket a granuláris szintézis bemutatásánál (5. fejezet) ismertettünk. A módszer két legfontosabb előnye a Fourier-analízissel szemben, hogy a frekvencia-felbontás logaritmikus (így minden oktávban egyenlő sűrűséggel kapunk összetevőket), továbbá nincs szükség a vizsgált jel „szeletelésére”: e módszer „beépítve” tartalmazza az időfelbontást. Ez feleslegessé teszi az ablakfüggvényeket is, így a végeredményt sem torzítja semmi. Az eljárás legnagyobb hátránya, hogy ez a „végeredmény” nem maga a spektrum, így a Fourier-analízis köré felépített digitális jelfeldolgozás világában (egyelőre) nem használható jól.

10.2. Megvalósítás

A 10.5. ábrán bemutatott `HApp_10_01` `(OSX)` `(WIN)` program segítségével kísérletezhetünk a Fourier-transzformációval. A felső sorban levő legördülő menüből választhatjuk ki az analizálni kívánt hangot. Ennek hullámforma-reprezentációját a legfelső hullámforma-kijelző mutatja. A „*Pozíció*” csúszka segítségével választhatjuk ki az analizálni kívánt jelszakaszt; a „*Lejátszás*” bekapcsolásakor a csúszka magától végigpásztázza a teljes hangot. Az analízis beállításait a jobb oldali mezők segítségével végezhetjük. Megjegyezzük, hogy a „*Fourier ablak mérete*” beállítás a már említett felülmintavételezést szabályozza, így ez nem lehet kisebb, mint a valódi ablakméret (ezt a program nem is engedné). A kiválasztott szakasz spektrumát a legalsó kijelzőn láthatjuk. Ennek skálázását (lineáris vagy decibel alapú) a megfelelő kapcsolóval változtathatjuk meg.



10.5. ábra. Hang le-ok Fourier-analízisét végző mintaprogram. Figyeljük meg, hogy az ablakfüggvények változásával hogyan változik az egyes csúcsok élessége, illetve a hozzáadott háttérzaj mértéke.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

A program képes az analízisadatok exportálására (a fázisok kivételével). Ezt az alsó sorban szereplő gombok segítségével tehetjük meg. A „*Frekvenciák*” és az „*Amplitúdók*” gomb megnyomásakor felugró ablakban az adatok „ömlesztve”, szóközzel elválasztva szerepelnek. Amennyiben mindent exportálunk, az ablakban egy kétoszlopos táblázat jelenik meg, melynek első oszlopában a frekvenciák, másodikban az amplitúdók szerepelnek. A frekvenciák mértékegysége mindig hertz; az amplitúdóké decibel vagy lineáris amplitúdó, a kijelző skálázását változtató kapcsoló állásának megfelelően.

A következő fejezetben látni fogjuk, hogy az analízis-adatok manipulálásán keresztül izgalmas effektusok érhetőek el. Az analízis-adatok tanulmányozásával jobban megérthetjük az ott bemutatandó eszközök működését.

10.3. Összegzés

A jelanalízis két célt szolgál. Az egyik, hogy kvantitatív információt nyerjünk a hangok belső szerkezetéről, elsősorban a hangszínt jellemző spektrum struktúrájáról. A másik, hogy ezt az információt felhasználva módosíthassuk a hangot (ennek részleteit a következő fejezet ismerteti). A legelterjedtebb módszer a Fourier-analízis, hiszen ez közvetlenül a szinuszos összetevőire bontja a jelet. Ám mivel a hang mikroszkopikus időskáláján az idő és a frekvencia összefügg, ezért a felbontás sosem lehet időben és frekvenciában egyszerre pontos: vagy az időfelbontás lesz túl szakaszos, vagy idegen frekvenciák (torzítások, zajok) jelennek meg a spektrumban. Különböző ablakfüggvények, illetve felülmintavételezés alkalmazásával a felbontás javítható, de „tökéletessé” nem tehető.

10.4. Feladatok

1. Vizsgáld meg, hogy különböző analízis-paraméterekkel milyennek látod a spektrumot! Vegyél fel egy hangot, és nyisd meg a `HApp_10_01` `(OSX)` `(WIN)` program segítségével. Próbálj ki legalább két ablakfüggvényt, mindegyiket legalább két eltérő ablakmérettel. Az így kipróbált 4 beállítást rendszerezd aszerint, hogy a kapott spektrum mennyire közelítette az eredeti hang (általad elképzelt) „valódi” spektrumát. Most vegyél fel egy új, az előzőtől lényegesen eltérő karakterű hangot, és készítsd el a spektrumait ugyanezzel a 4 beállítással! Rendezd újra sorba a beállításokat, de most a második hang szerint. Megmaradt a beállítások sorrendje?

2. Hangmagasság meghatározása. Vegyél fel egy harmonikus hangot, aminek ismered a hangmagasságát, és nyisd meg a `HApp_10_01` `(OSX)` `(WIN)` program segítségével! A megfelelő analízis-paraméterek megtalálása után exportáld az adatokat (a frekvenciákat és az amplitúdókat egyaránt). A 0 Hz-től felfelé indulva keresd meg az első öt olyan frekvenciát, ahol az amplitúdó-értékekben ún. *lokális maximumot*, „csúcsot” találsz („lokális maximumnak” azt a frekvencia-amplitúdó párost értjük, amelyre teljesül, hogy a hozzá közel eső frekvenciáknak az amplitúdói mind kisebbek, mint a lokális maximumé[3]). Hogyan viszonyulnak az így megtalált frekvencia-értékek az eredetileg felvett hang alapfrekvenciájához? Mi lehet az oka annak, ha az így kapott frekvenciák nem az alaphang felhangjai? Mi lehet az oka, ha azok?

[3] Pl. a $3,5,4,8,6,3$ sorozatban a sorozat 2. és 4. elemei lokális maximumok, melyek értékei 5, illetve 8.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

3. Ismételd meg az előző feladatot inharmonikus hangokkal és zajokkal is! Hogyan lehet ránézésre különbséget tenni egy harmonikus, egy inharmonikus és egy zajos hang spektruma között?

© *Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola*

Open Access 2026-11-02

11. fejezet

Reszintézis és alkalmazásai: transzpozíció, nyújtás, szűrés, keresztszintézis

A 10. fejezetben láttuk, hogy hogyan bonthatunk egy jelet spektrális komponenseire. E fejezetben azt mutatjuk meg, hogy az így nyert adatokból hogyan állíthatóak elő új hangok.

A spektrum alapján történő szintézisen túl pár gyakorlati példát is ismertetünk, amelyek mind az analízis–spektrummódosítás–reszintézis sémát követik: a *transzponálást*, a *nyújtást*, a *spektrális szűrést* és a *keresztszintézist*.

11.1. Elmélet

11.1.1. Reszintézis

Inverz Fourier-transzformáció

A Fourier-transzformáció *invertálható*: egy adott jel spektrumából az eredeti jelet (matematikai úton) egyértelműen ki lehet nyerni. Az inverz eljárás pontosan ezen alapul: az egyes ablakok adataiból matematikai úton kinyerjük azt a hullámformát, amit ezen ablakok spektrumai leírnak, majd összeadjuk őket.

A módszer csak akkor adja vissza tökéletesen az eredeti jelet, ha az analízis során a jelből kivágott „szeletek” átfedés nélkül, ám hézagmentesen illeszkedtek egymás mellé, továbbá négyszögletes ablakfüggvényt alkalmaztunk, felülmintavételezés nélkül (e fogalmak magyarázatát lásd a 10. fejezetben). Ha nem négyszögletes ablakfüggvénnyel dolgoztunk, illetve felülmintavételeztük a mintát, az analízis által keltett torzulások megjelennek a szintetizált végeredményben. Emellett, ha az analízis során kivágott jelszeletek között átfedés volt, az ablakok inverz Fourier-transzformálásakor kapott szeletek is át fognak fedni, ami további torzulást (érdességet) okoz. Ezzel szemben az analízis szempontjából a fenti feltételek (négyszögletes ablak, nincs felülmintavételezés, az analizált szakaszok nem fednek át) általában közel sem optimálisak. Megjegyezzük, hogy bizonyos feltételek teljesülése esetén (pl. Hann-ablak, 50%-os átfedéssel az egyes analízis-szakaszok között) az eredmény meglepően jó, ezért a gyakorlatban sokszor ilyen (vagy hasonló) beállításokkal alkalmazzák ezt az eljárást.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

Additív és szubtraktív módszerek

Az additív és a szubtraktív szintézis során a cél egy adott spektrum felépítése (az additív módszer a „nullából építkezve” rakja össze a spektrumot, míg a szubtraktív a „zajt kifaragva” jut el a célig). Mindkét módszer használható a reszintézishez.

Általában akkor alkalmazzuk e megoldásokat, ha a spektrumot előzetesen egyszerűsíteni tudjuk: ki tudunk választani kiugró frekvenciákat és/vagy formáns (esetleg zajos) területeket a spektrumból, melyeket egy-egy paraméterpárral helyettesíthetünk. Az additív hozzáállás esetén az így nyert amplitúdó- és fázisadatokat egy (kellően sok csatornás) additív szintetizátor bemenetére kötjük;[1] a szintetizált jel így pontosan azokat a frekvenciákat (és olyan amplitúdókkal) fogja tartalmazni, amelyek a létrehozandó spektrum legfontosabb komponensei. A szubtraktív megoldás (a „*vokóder*”) az analízisből nyert formánsok és egyéb zajos területek alapján egy sokcsatornás ekvalizátort vezérel, melyet egy sűrű spektrumú forrásra (általában fehérzajra vagy impulzussorozatra) csatlakoztatunk; ekkor tehát a célspektrumot e forrás szűrésével érjük el.

Felépítése miatt a vokóder kiválóan alkalmas bizonyos spektrális manipulációk (pl. keresztszintézis) megvalósítására, amennyiben nem fehérzajt, hanem a manipulálandó jelet használjuk forrásként. A legtipikusabb példa erre az élő beszéd manipulálása (gondoljunk a 70-es vagy a 80-as években forgatott sci-fik mesterséges hangon beszélő robotjaira.)

Hibrid eljárások

Akár az oszcillátorbankot, akár a vokódert véve alapul, nyilvánvalónak tűnik, hogy minél több csatornával dolgozunk, annál pontosabban tudjuk előállítani a megkövetelt spektrumot. Más szóval, minél több szintézisegységet használunk, annál kisebb lesz a különbség ugyanazon spektrum oszcillátorokkal, illetve vokóderrel létrehozott példányai között.

A legjobb példa erre a „*fázisvokóder*”, melyet egyaránt felfoghatjuk vokódernek, oszcillátorbanknak, vagy akár egy speciális inverz Fourier-transzormátornak is. A fázisvokóder az egyes ablakokból kinyert adatok alapján becslést ad a spektrum komponenseinek amplitúdó- és fázis-burkológörbéire, majd az így nyert trajektóriák segítségével „korrigálja” a nyers adatokat. Egyszerűbb esetekben ez a korrekció egy interpolációt jelent, mellyel az analízisből nyert amplitúdó- és fázisértékekből egy sűrűbb spektrogramot készítünk. Ezt a spektrogramot (megfelelő paraméterek választásával) inverz Fourier-transzformálva jobb minőségű hangot nyerünk, mint a „naív” inverz Fourier-módszerrel. A bonyolultabb fázisvokóderek ezen felül a nyers adatok megfelelő csoportosítása, szűrése és manipulálása útján érnek el tisztább hangzásokat.

Ahogy említettük, a fázisvokóderek gyakran inverz Fourier-transzformáción alapulnak, de egyes implementációk a kiszámított trajektóriák alapján egy kiterjedt oszcillátorbankot (vagy akár egy sokcsatornás „hagyományos vokódert”) hajtanak meg.

11.1.2. Nyújtás

Ha egy felvett hang lejátszási sebességét változtatjuk, az eredmény nemcsak gyorsabb, vagy lassabb lesz, de hangszíne is megváltozik: az eredeti hang transzponálódik, illusztrálva, hogy az idő és a frekvencia nem független (lásd a 11.1. hangpéldát). Ahhoz, hogy e két jelenséget

[1] Az ilyen, valós időben vezérelt sokcsatornás additív szintetizátort gyakran „*oszcillátorbanknak*” hívjuk.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

(transzpozíció, illetve nyújtás) egymástól függetlenül kezelhessük, bonyolultabb műveletekre van szükség.

A kiindulási pont mindkét esetben az eredeti hang spektrális felbontása. Az ablakolt spektrumok megfelelő módosítása után, reszintézis segítségével állítjuk elő a transzformált hangot.

Nyújtáskor az egyes ablakolt spektrumokat egyáltalán nem módosítjuk, pusztán az ablakok időtartamait változtatjuk meg. Például, ha az eredeti hang felbontásakor 50 ms-os ablakokat alkalmaztunk, az időtartam megduplázásához az szükséges, hogy a reszintézis során 100 ms-osnak tekintsük ezeket az ablakokat. Mivel a spektrumot nem kell módosítani, ez az eljárás lényegesen egyszerűbb, mint a transzpozíció.

Mivel a frekvencia-, illetve időfelbontás érzékeny a választott ablak méretére és fajtájára, különböző hangokat más-más paraméterekkel célszerű nyújtani: ütött, zajos hangokhoz általában kisebb, míg harmonikusakhoz inkább nagyobb ablakméretet szokás használni. A 11.2. és a 11.3. hangpéldák e két típust illusztrálják. A beszédhang nyújtása általában nagyobb körültekintést igényel, mivel itt a hang rövid változásai és a harmonikus tartalmuk egyaránt fontosak (lásd a 11.4. hangpéldát).

Megjegyezzük, hogy a nyújtás más módszerekkel is megvalósítható, ám ezek többnyire rosszabb hatásfokúak, mint az itt ismertetett eljárás.

11.1.3. Transzponálás

A transzpozíciót többféle módon is megvalósíthatjuk. Az egyszerűbb megoldás az, ha (a fentebb leírt jelenséget kihasználva) először megváltoztatjuk a minta lejátszási sebességét (így elérjük a kívánt transzpozíciót), majd megfelelő mértékű nyújtással visszaállítjuk az eredeti hang hosszát.

Sajnos az így elérhető transzponálás mesterségesnek hat. Ennek az az oka, hogy a legtöbb természetes hang spektruma számos olyan formánst tartalmaz, amelyek helye független az adott hang magasságától. A 7. fejezetben ismertettük az emberi hangképzést; amint láttuk, az énekelt hang magasságát egyedül a hangszálak rezgési frekvenciája határozza meg, a formánsok minden esetben ugyanott vannak. Emiatt pontosabb eredményt kapunk, ha a következőképpen járunk el.

Először minden ablakolt spektrumban meghatározzuk e spektrumok burkológörbéit. Ennek az az értelme, hogy ezek a burkológörbék tartalmazzák az adott spektrum formánsait. Második lépésben a spektrum frekvenciáit átskálázzuk (ez maga a transzpozíció). A harmadik lépésben az átskálázott[2] spektrumot moduláljuk az eredeti burkológörbével, így a transzponált hangzásba visszajuttatjuk az eredeti formánsokat. E három lépést a 11.1. ábra illusztrálja. Végül a kapott ablakolt spektrumokkal elvégezzük a reszintézist.

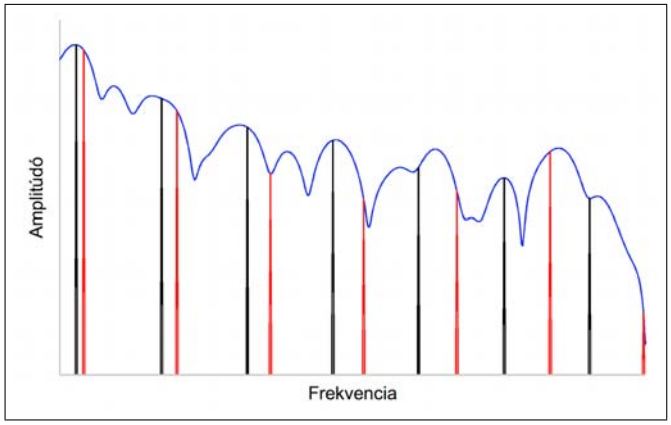
Fontos megjegyezni, hogy minél jobban nyújtjuk, illetve transzponáljuk a hangot, annál „messzebb” jutunk az eredeti spektrumtól, vagyis annál jobban felerősítjük a Fourier-analízis során kapott torzításokat. Emiatt a hangszín csak a kis változtatások során marad kellően közel az eredetihez (mindkét eljárás esetén).

A legtöbb gyakorlati alkalmazás nemcsak konstans transzpozíciót és nyújtást enged; akkor válnak igazán érdekessé e módszerek, amikor időben változó mértékben alkalmazzuk őket (pl. glissando, ritenuto stb.).

[2] Megjegyezzük, hogy az átskálázás előtt az eredeti spektrumot ki is kell simítani – ezt terjedelmi okokból nem fejtjük ki.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



11.1. ábra. Transzpozíció az eredeti spektrális burkológörbe megtartásával. Az eredeti spektrumot a fekete, a spektrális burkológörbét a kék, az új spektrumot a piros vonalak jelzik. Figyeljük meg, hogy (mivel transzponáláskor a frekvenciák aránya *az, ami változatlanul marad) a távolság az új és az eredeti komponensek között a frekvencia növekedésével arányosan nő.*

A 11.5. és a 11.6. hangpéldák ugyanazt a szöveget transzponálják. A transzpozíció mindkét esetben egy oktávval mélyebbről indul, majd egy-egy glissando segítségével egy oktávval magasabban ér véget, mint az eredeti jel. A két példa közötti különbség, hogy míg a 11.5. példa a teljes jelet transzponálja, a 11.6. példa a fent leírtak szerint megtartja az eredeti hang spektrális burkológörbéjét. A 11.7. és a 11.8. hangpéldák ugyanezt a jelenséget illusztrálják, de az 1.5. (zenekari) hangpéldából kiindulva.

11.1.4. Spektrális szűrés

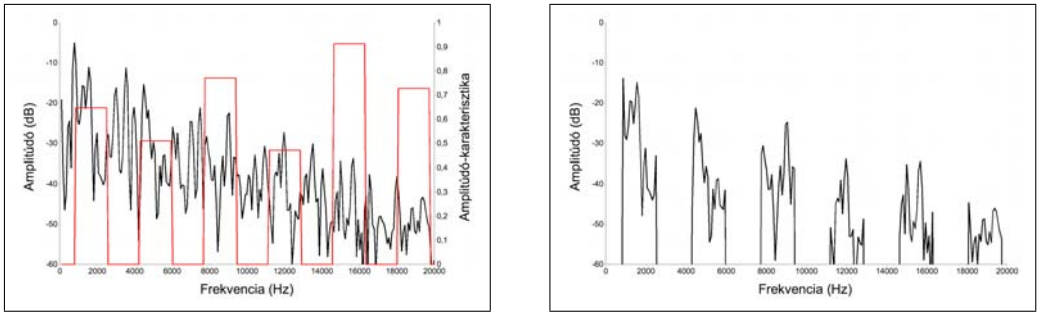
A legaprólékosabb szűrési módszerek egyike, melyet a 11.2. ábra illusztrál. Alkalmazása során az analízisből nyert ablakolt adatokon akár ablakonként végighaladva, egyesével változtathatjuk meg a spektrális struktúrát. Ráadásul az egyes frekvenciák amplitúdóit akár egyesével szerkeszthetjük, így szabadon szerkeszthetővé válik a szűrő amplitúdó-válasza. A szűrési paraméterek egész komplexek lehetnek: lehetséges akár egy kép alapján is származtatni őket, amint az a 11.3. ábrán látható. Habár tökéletes módszernek tűnik, nem szabad elfelejtenünk, hogy az analízis és a reszintézis során a jelhez mindenképp hozzáadódó torzulások miatt csak akkor érdemes a módszerrel élni, ha az alkalmazott szűrőprofil komplexitása „elnyomja” e torzulásokat. Egyszerű karakterisztikáknál (sáváteresztő, aluláteresztő stb.) általában nem ez a helyzet.

11.1.5. Keresztszintézis

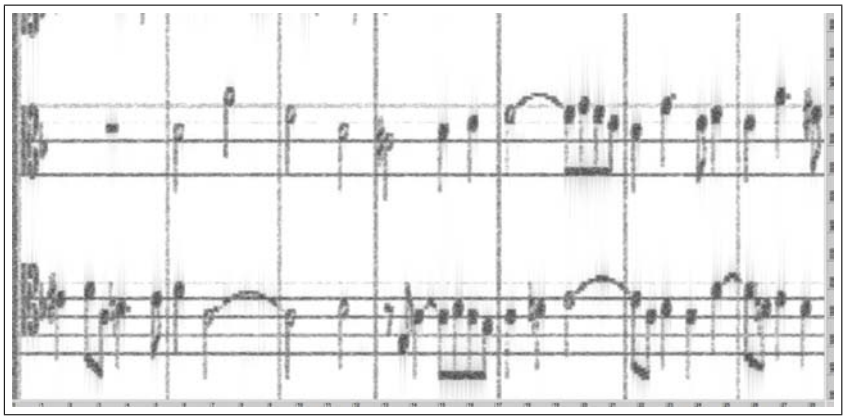
A módszer egyfajta „spektrális modulálásként” fogható fel, és mint ilyen, két független jelre épül: egy vivőjelre (ez általában statikus és a spektruma sűrű, pl. szél) és egy modulálójelre (mely általában gyorsan változó, pl. emberi beszéd). Az eljárás során mindkét hangot analizáljuk (pl. ablakolt Fourier-transzformációval), majd a modulálójelből kapott ablakok spektrumainak kiszámítjuk a burkológörbéit. E burkológörbékkel azokat a spektrumokat „moduláljuk”, melyeket a vivőjel ablakolt analíziséből nyerünk. Az így kapott spektrális adatokból reszintézis útján új hangot állítunk elő. Az eljárást a 11.4. ábra szemlélteti.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



11.2. ábra. Többsávos spektrális szűrő. Bal oldalon a forrás spektruma és a szűrő-karakterisztika (piros vonal), jobb oldalon a szűrt spektrum látható. A 11.9. hangpélda többsávos szűrővel szűrt fehérzajt tartalmaz.



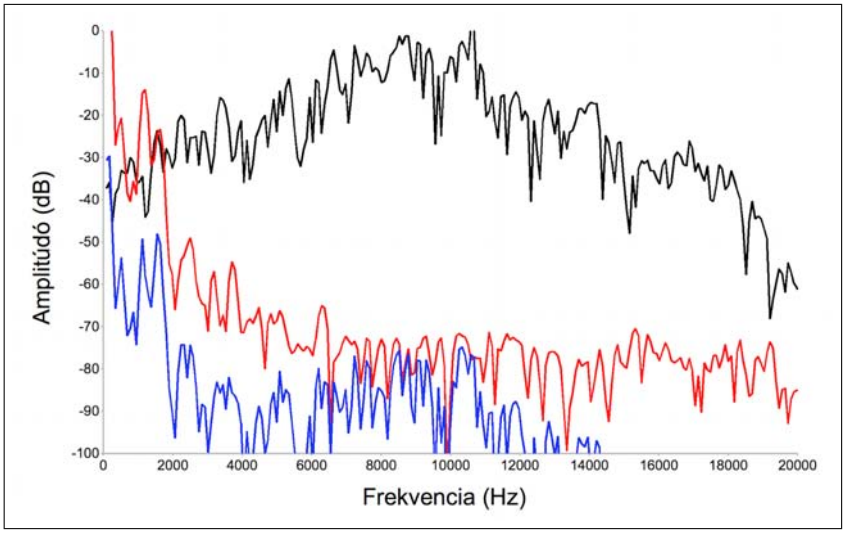
11.3. ábra. Kép alapján vezérelt komplex spektrális szűrés. Az eljárás során a képet alkotó képpontokat szűrő-paraméterekként értelmezzük. Először vízszintesen felosztjuk a képet annyi szeletre, ahány FFT ablakunk van; ezután az egyes szeletekben levő képpontokat hozzárendeljük a nekik megfelelő FFT-ablakoknak. Egy-egy ablakon belül úgy alakítjuk ki a szűrőpro lt, hogy a képpontokhoz függőleges elhelyezkedésük alapján frekvenciaértékeket rendelünk, és az így meghatározott frekvenciákon olyan mértékű szűrést írunk elő, mint amilyen az adott frekvenciának megfelelő képpont intenzitása (az intenzitás annak a mértéke, hogy az adott pont mennyire sötét vagy világos). A 11.10. hangpélda elkészítéséhez fehérzajt vettünk alapul, melyet egy fényképpel szűrtünk. Az ábra az így kapott hang szonogramjának egy (1 200 Hz és 4 000 Hz közötti) részletét mutatja.

A módszerrel olyan hangot kapunk, melynek hangszínét ugyan a vivőjel adja, de a „karaktere” a modulálójeltől származik. Egyszerű példa a „beszélő szél”, vagy a „beszélő zenekar”, amikor a szél, illetve a zenekar „emberi nyelven szól” (lásd a 11.11. hangpéldát).

Megjegyezzük, hogy a „keresztszintézis” egy tágabb kifejezés, melyet nagyjából minden olyan eljárásra szoktak alkalmazni, ahol két jel modulálja egymást a spektrumaikon keresztül; a fentebb bemutatott módszer tehát csak egy bizonyos fajta keresztszintézist takar.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



11.4. ábra. Keresztszintézisben részt vevő spektrumok. A fekete spektrum a vivőjelhez tartozik (zajszerű forrás), a piros pedig a moduláló jel (hangszerhang). A végeredményt a kék spektrum mutatja.

11.2. Megvalósítás

A `HApp_11_01` `(OSX)` `(WIN)` és a `HApp_11_02` `(OSX)` `(WIN)` programok egyszerűen mutatják be a lejátszási sebesség megváltoztatásán alapuló nyújtást (lásd a 11.5. ábrát). A hangfile kiválasztása után akár a csúszka, akár a számdoboz segítségével állíthatunk a lejátszás sebességén (a normális lejátszás az egységnyi sebességnek felel meg). Megfigyelhető, hogy a lejátszás sebességének változtatásával a transzpozíció is változik. Habár nem alkalmas sem igazi transzponálásra, sem nyújtásra, az eljárás egyszerű kivitelezhetősége és könnyű kezelhetősége miatt mégis népszerű; a (főleg) DJ-k által kedvelt „*scratching*” technika (lásd a 11.12. hangpéldát) ezen az elven alapul.

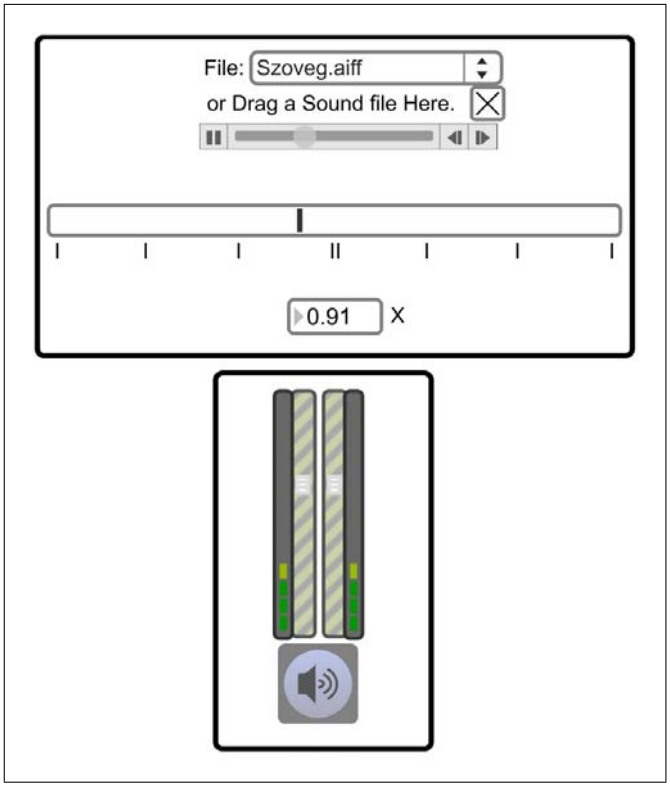
A `HApp_11_03` `(OSX)` `(WIN)` ezzel szemben olyan program, ami az analízis-reszintézis segítségével végzi a transzponálást. A nyújtást továbbra is a lejátszási sebesség változtatásával éri el; ám mivel a transzponálást a nyújtástól függetlenül tudjuk szabályozni, ezért tetszőleges nyújtási-transzponálási kombinációt lehetséges létrehozni. A példa tartalmaz egy spektrális szűrőt, valamint egy amplitúdó-modulátort is. A kezelőfelületet a 11.6. ábra mutatja be.

Az analízis és a reszintézis paramétereit (pl. analízisablak mérete és típusa stb.) a program automatikusan állítja be. A transzpozíció és a nyújtás külön-külön kezelésével hallhatóvá válik, hogy az analízis során fellépő torzulások (bővebben lásd a 10. fejezetet) hogyan befolyásolják a végeredményt. Különösen akkor válnak e torzulások kézzelfoghatóvá, ha a lejátszást nagyon lelassítjuk. Extrém esetben az is előfordulhat, hogy a végeredményben már semmilyen nyomát sem leljük az eredeti jelnek, annyira dominánssá válnak a torzulások.

A keresztszintézist a `HApp_11_04` `(OSX)` `(WIN)` program valósítja meg (lásd a 11.7. ábrát). A két forrásjelet a képernyő tetején levő két hangforrás-modul segítségével adhatjuk meg; a középen levő csúszkák („*Amplitúdó*” és „*Fázis (Frekvencia)*”) azt határozzák meg, hogy a létrehozott jel spektrumának amplitúdóira, illetve fázisaira (utóbbi gyakorlatilag a frekvenciákért felel) mekkora hatással legyen a bal, illetve jobb oldali forrás.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



11.5. ábra. Lejátszás sebességének változtatása Max programnyelven. A `HApp_11_01` `(OSX)` `(WIN)` és a `HApp_11_02` `(OSX)` `(WIN)` példák kezelőfelületei ugyanúgy néznek ki, ám míg előbbi esetén a lejátszás sebessége tetszőlegesen állítható, utóbbinál a sebesség a csúszka elengedése után visszatér a „normálishoz” (mint egy valódi lemezjátszón).

11.3. Összegzés

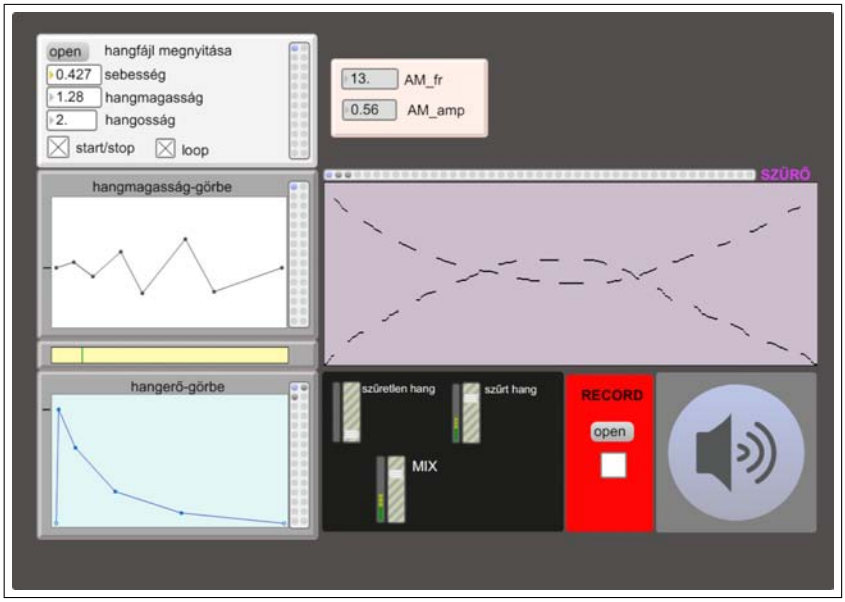
Hangszín. Az analízissel és reszintézissel létrehozható hangok elsősorban az eredeti, analizált hangtól függnek. A technikát gyakran használják transzponálásra és nyújtásra, de számos egyéb reszintézisre épülő módszer is létezik, mint pl. a spektrális szűrés vagy a keresztszintézis.

Paraméterek. Attól függően, hogy a spektrumot milyen mértékben módosítjuk, előtérbe kerülhetnek az analízist és a reszintézist meghatározó paraméterek (pl. használt ablakfüggvények, Fourier-ablak mérete stb.). A spektrum drasztikus módosításakor ezek a paraméterek – az általuk keltett, eltérő jellegű torzulásoknak köszönhetően – nagymértékben befolyásolják a kialakuló hangszínt; kisebb módosítások esetén nem jutnak lényeges szerephez.

Műveletigény. A módszerek számításigénye nagyban függ az elvárt tisztaságtól. A Fourier-transzformáció – amennyiben az ablakban levő jelminták száma valamilyen 2-hatvány – nagyon gyorsan elvégezhető, ám a spektrum megfelelő transzformálása a legtöbb esetben erőforrásigényes. Kiváltképp igaz ez a fázisvokódert alkalmazó reszintézis esetén, amikor a spektrumon végzett műveleteknek az egyes frekvenciasávok fázisaival is számolniuk kell.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



11.6. ábra. Spektrális analízis-reszintézis útján végzett műveletek. A jobb oldali ekvalizátorral vezérelhetjük a spektrális szűrést, míg a bal oldalon a hang függetlenül transzponálható és nyújtható. A megfelelő számdobozban megadott hangmagasságszorzót a hangmagasság-görbe segítségével tudjuk modulálni. A jobb oldali felső egység egyszerű amplitúdó-modulációt tesz lehetővé.

11.4. Feladatok

1. Készítsd el ugyanannak a hangnak három változatát. Az első változat legyen egy oktávval mélyebb (de változatlan időtartamú), a második legyen kétszer olyan hosszú (de változatlan hangmagasságú), a harmadik legyen egy oktávval mélyebb és kétszer olyan hosszú, mint az eredeti hang. Hasonlítsd össze a három hangot hangminőség és az elért hatás tekintetében.

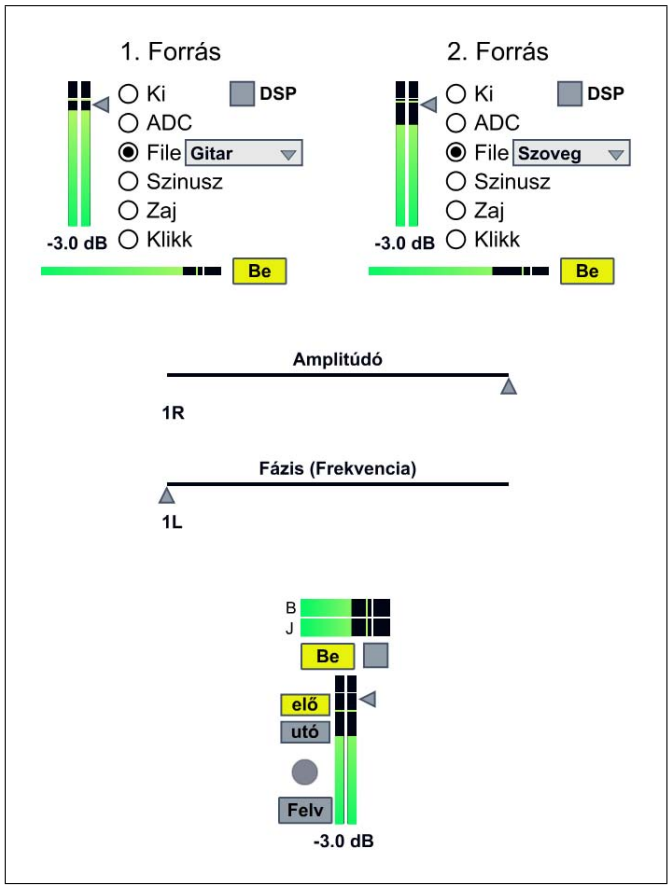
2. Hasonlítsd össze a spektrális szűrést a „hagyományos” szűrési eljárásokkal! A `HApp_11_03` `(OSX)` `(WIN)` program szűrőjével rajzold meg egyesével a 4.2. ábrán látható karakterisztikákat, és hallgasd meg, hogy ezek hogyan módosítják a beérkező hangokat. Miben térnek el az így szűrt hangok a 4. fejezetben bemutatott szűrőkkel szűrt hangoktól?

3. A `HApp_11_03` `(OSX)` `(WIN)` program szűrőpanelének segítségével rajzolj egy formánsokkal tarkított, harmonikus spektrumot! Olvass be egy hangot és transzponáld! Az eredményt vezesd át a szűrőn! Mit tapasztalsz?

4. A `HApp_11_04` `(OSX)` `(WIN)` egyik forrása legyen a mikrofon, a másik forrásnak válassz egy harmonikusan gazdag, ám nem zajszerű forrást (pl. kórus, zenekar, szólóhangszer). Állítsd be úgy a programot, hogy a szintetizált hang amplitúdóit a mikrofon, frekvenciáit a másik forrás vezérelje. Mit tapasztalsz, ha a mikrofonba beszélsz? Cseréld fel a két forrást! Most mit tapasztalsz? Kísérletezz a frekvenciák és amplitúdók különböző arányaival a két forrás között!

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



11.7. ábra. Keresztszintézis. *A kimeneti jel amplitúdóit az* $A_0[n] = A_1[n]\cos\left(\frac{\pi}{2}x\right) + A_2[n]\sin\left(\frac{\pi}{2}x\right)$, *míg fázisait a* $\phi_0[n] = \phi_1[n]y + \phi_2[n](1-y)$ *képlet alapján határozhatjuk meg (*x *és* y *az „Amplitúdó” és „Fázis (Frekvencia)” csúszkák értékei, melyek 0 és 1 között lehetnek).*

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

12. fejezet

Jelfeldolgozás I: késleltetésen alapuló eszközök

Ebben a fejezetben késleltetésre, illetve szűrésre épülő eszközöket mutatunk be. A késleltetés, illetve szűrés fontos közös jellemzője, hogy – amennyiben paramétereiket időben nem változtatjuk – nem hoznak létre új frekvenciákat a spektrumban. Az ilyen eszközöket (vagyis amelyek nem adnak új frekvenciákat a spektrumhoz) *lineárisnak* nevezzük.

12.1. Elmélet

12.1.1. Szűrés vagy késleltetés?

Gondoljuk meg, mi történik, ha egy hangot nagyon rövid idővel késleltetünk, majd összekeverjük a késés nélküli eredetijével: azok a frekvenciák, amelyek periódusideje megegyezik a késleltetés idejével, felerősödnek, míg azok, amelyek pont ellentétes fázisban találkoznak (vagyis a periódusidő pont a késési idő duplája), kioltják egymást. Mivel a szűrők pont ezt használják ki, ezért a szűrés és a késleltetés ugyanannak a jelenségnek a különböző megnyilvánulásai: „szűrésről” beszélünk, ha a spektrum jelentősen megváltozik (ekkor a jel késése nagyon kicsi), míg „késleltetésről”, ha a késés jelentős (ekkor a hangszín változása elhanyagolható).

Egy lineáris eszköz általános leírásához a bejövő hang spektrumában szereplő amplitúdókra, illetve fázisokra (lásd 10. fejezet) gyakorolt hatását kell megadnunk. E „fázisokra gyakorolt hatást” *fázistolásnak*, vagy *fáziskésésnek* hívjuk: azt fejezi ki, hogy a jelet felépítő (különböző frekvenciájú) szinuszhullámok külön-külön mennyit késnek. Habár léteznek olyan eszközök, ahol az egyes frekvenciákhoz tartozó fázistolások igen eltérőek lehetnek, miközben az egyes amplitúdók változása is igen jelentős (a jelenséget *spektrális késleltetésnek* hívjuk), a legtöbb eljárás során a fázistolás minden frekvencián azonos nagyságrendbe esik. Ezt az átlagos fázistolást nevezzük *késleltetési időnek*.

A késleltetési idő szerint három nagy csoportra oszthatjuk e módszereket:

0–10 ms: Ezek a köznapi értelemben vett *szűrők*. Az ilyen rendszereknél az amplitúdóváltozás erősen frekvenciafüggő. A 12.1. hangpéldában fehérzajt késleltettünk 4 ms-mal, az így kapott jelet hozzáadtuk az eredeti zajhoz.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

10–50 ms: Ez a késleltetés már túl hosszú ahhoz, hogy a frekvenciákat jelentősen befolyásolja, de túl rövid, hogy a késleltetett jelet külön észlelni tudjuk. Általában a hang „teresítését” szolgáló modulokra jellemző ez a késleltetési időtartomány. A 12.2. hangpélda 30 ms-mal késleltetett zongorahangot tartalmaz.

50 ms felett: A késleltetett hang egyértelműen elkülönül az eredetitől, visszhangot képezve, miközben a spektrum gyakorlatilag semennyit sem módosul. Ezek a köznapi értelemben vett *késleltetők*. A 12.3. hangpélda egy rövid jelet és annak 300 ms-mal késleltetett változatát mutatja be.

A lineáris rendszerek fontos eleme a *visszacsatolás*, melynek során a végeredményként kapott hangot visszakötjük a rendszer bemenetére. Ezzel „kihangsúlyozható” a karakterisztikus viselkedés. Szűrők esetén például a levágási frekvenciánál tapasztalható levágás tehető meredekebbé, visszhangok képzésekor pedig egyetlen visszhang helyett egy teljes sorozatot kapunk, ami valósághűbbé teszi a hatást. Visszacsatolás előtt módosíthatjuk a visszacsatolt jel amplitúdóját. Ennek függvényében a következő eredményekre számíthatunk:

Halkítás: Stabil hangzás (lásd a 12.4. hangpéldát).

Változatlan amplitúdó: Amíg van beérkező jel, az eredmény folyamatosan hangosabbá és telítettebbé válik. Ha egy ponton kikapcsoljuk a bejövő hangot, egy szabályosan ismétlődő hangmintát, loopot kapunk (lásd a 12.5. hangpéldát).

Erősítés: Az eszköz megállás nélkül hangosodik, „*gerjedés*” jön létre (lásd a 12.6. hangpéldát).

A „klasszikus szűrők” tulajdonságait a 4. fejezetben ismertettük. A továbbiakban néhány speciális eszközt mutatunk be, melyek késleltetésre, illetve szűrésre épülnek.

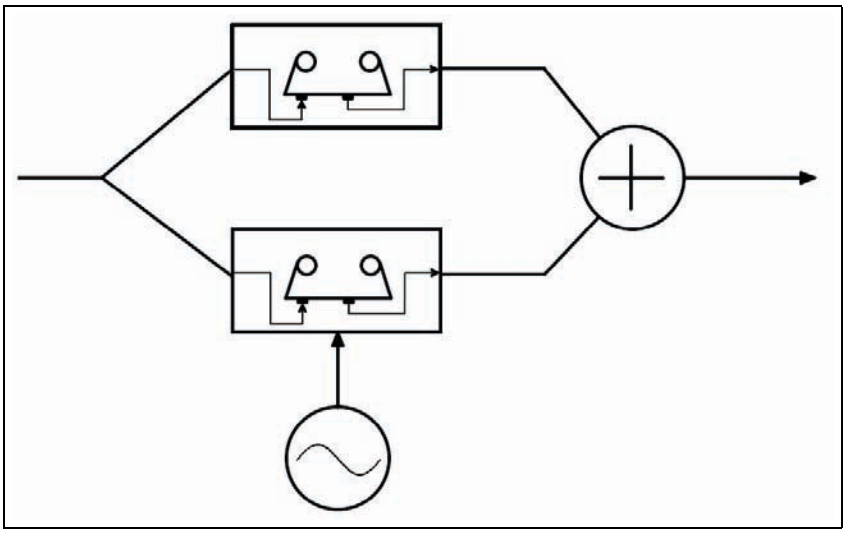
12.1.2. Flanger, fázistoló

A *angerben* a beérkező jelet két párhuzamos ágra szedjük, mindkét ágat késleltetjük. Az egyik ág konstans időt késik, a másik késleltetését folyamatosan változtatjuk, de úgy, hogy átlagosan ugyanannyi legyen a két jel késése. A folyamatot a 12.1. ábra mutatja. A két ág összekeverése után azt tapasztaljuk, hogy az eredeti jel spektrumához képest a végeredményben egyes frekvenciák eltűnnek, sőt, e „lyukak” egyenlő távolságra helyezkednek el egymástól (vagyis egy „harmonikus űr” keletkezik). Ráadásul a második ágon folyamatosan változó késési idő miatt ezek a lyukak fel-le vándorolnak a spektrumban. A 12.7. hangpélda fehérzajjal mutatja be a flangert.

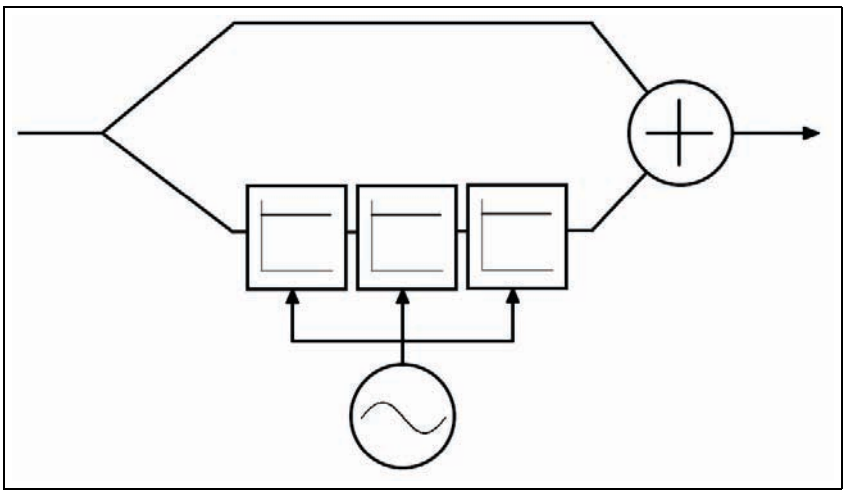
A *fázistoló* ezzel szemben *mindent áteresztő szűrőkből* (lásd a 4.2. ábrát) építhető fel a legkönnyebben, ahogy a 12.2. ábrán látható. Ezek olyan szűrők, amelyeknek a fázistolása nagyon kicsi (ezért szűrőnek tekintjük), viszont a spektrum amplitúdóit egyáltalán nem módosítják; így egyedül a fázisok azok, amelyek módosulnak egy ilyen eszközben. Ha tehát az eredeti jelet mindent áteresztő szűrők rendszerén engedjük át, az egyes frekvenciák fázisait különböző mértékben tudjuk eltolni; ha az így kapott hangot összekeverjük az eredetivel, bizonyos frekvenciák kiemelődnek, míg mások gyengítődnek. A fázistolások mértékének (vagyis a mindent áteresztő szűrők paramétereinek) folyamatos változtatásával e kiemelések és gyengítések helye folyamatosan változtatható. A 12.8. hangpélda fehérzaj segítségével illusztrálja a fázistoló működését.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



12.1. ábra. Flanger folyamatábrája. A felső késleltető konstans időt késik, az alsó késési idejét egy szinuszos oszcillátor modulálja.



12.2. ábra. Három szűrőt tartalmazó fázistoló folyamatábrája. A szűrők „középfrekvenciáit” (mely jelen esetben azt a frekvenciát jelenti, ami felett a szűrő megfordítja a fázist) egy szinuszos oszcillátor modulálja.

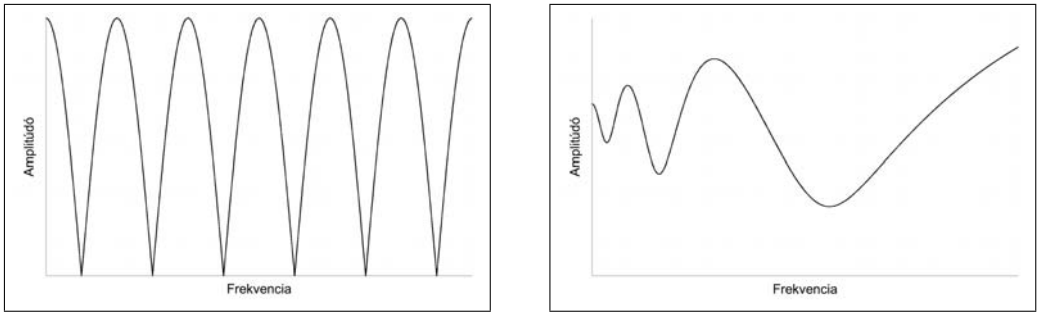
A fázistoló és a flanger hangja erősen hasonlít egymáshoz: a végeredmény mindkét esetben azt az érzetet kelti, mintha az eredeti hangszín „vándorolna”, miközben a hang fő paraméterei (hangmagasság, hangerő, dinamika stb.) változatlanok. A két módszer közötti fő eltérés, hogy míg a flanger spektrumában teljes kioltások keletkeznek, és ezek szigorúan harmonikus rendben követik egymást, a fázistoló erősítési és kioltási helyei, valamint ezek mélysége szabályozható (lásd a 12.3. ábrát). Ennek a szabályozhatóságnak az ára természetesen a jóval bonyolultabb paraméterezés.

12.1.3. Elhangolás, vibrato

A késleltetés változtatásával a bejövő jel elhangolódik. Ezt a hatást *Doppler-effektusnak* hívjuk; a természetben is könnyen megfigyelhető (pl. amíg egy mentőautó közelít felénk,

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



12.3. ábra. A anger (bal) és a fázistoló (jobb) jellemző spektrumai.

a sziréna hangja magasabb, míg távolodás közben mélyebb, mint az álló mentőautóé). Ha tehát növeljük a késési időt, a hangot mélyíteni tudjuk, a késési idő csökkentésével pedig magasítani (értelemszerűen utóbbit csak addig tehetjük meg, amíg a késési időt nullára nem csökkentettük). A mélyülés, illetve magasodás a késleltetési idő változásának pillanatnyi sebességével arányos, vagyis minél gyorsabban növeljük/csökkentjük a késleltetési időt, annál nagyobb transzpozíciót kapunk. A jelenség egy az egyben ugyanaz, mint a 11. fejezetben leírt, a visszajátszási idő manipulálásán alapuló transzponálási módszer. Emiatt az így létrehozott transzpozícióval a hang időtartama is változik.

Mivel a transzpozíció mértéke a késési idő megváltoztatásának a *sebességétől* függ, ezért ennek a „változási sebességnek" a folyamatos változtatásával a transzpozíció mértéke is folyamatosan változni fog. Ha tehát a késleltetési időt periodikusan változtatjuk (például egy szinuszos modulációval), vibrato érhető el (lásd a 12.9. hangpéldát). A pillanatnyi elhangolás a moduláló jel pillanatnyi sebességétől függ. Szinuszos modulálás esetén e sebesség legnagyobb értéke $2\pi A f_{\text{mod}}$, ahol A a moduláló jel amplitúdója, f_{mod} a frekvenciája.

12.1.4. Kórus

Egy valódi kórusban nincs két teljesen egyforma hang. Hiába énekelnek egy szólamot többen is, az egyes előadók eltérő hangszíne, intonálása, hangereje és reakcióideje miatt a szólam minden tagja kicsit eltérő módon járul hozzá a végeredményhez. Ezek az apró eltérések adják az egyes szólamok telt, gazdag hangzását.

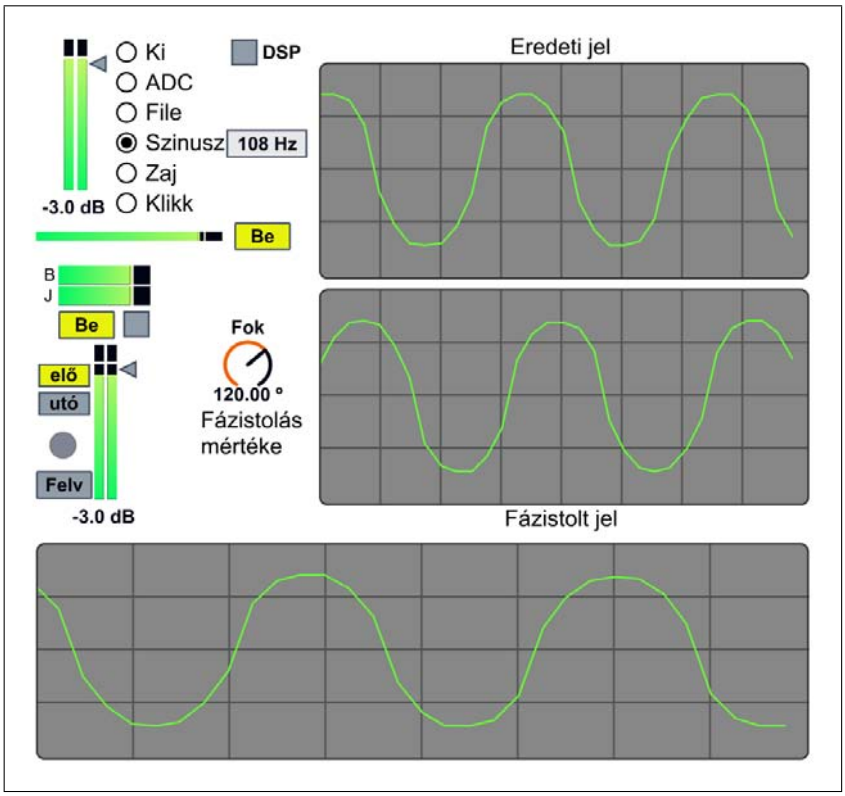
A kórus-effektek ezt a hatást utánozzák. Nincs általános formula, amivel egy valódi kórust szimulálni lehet, de az eszközök többsége a fenti módszereket kombinálja: a beérkező jelet több, önálló ágra bontja, majd az ágakat különböző mértékben modulálja, transzponálja, illetve eltérő fázistolásokat, szűréseket alkalmaz rajtuk. A végeredmény az egyes ágak eltérő arányú keverésével születik. A 12.10. hangpélda a következő szakaszban bemutatott kórust alkalmazza J. S. Bach BWV 1011-es csellószvitjének egy rövid részletére. Természetesen nagyon körültekintően kell egy kórus-effektet megépíteni: amennyiben túl kevés, vagy túl „szabályosan" paraméterezett eljárást használunk, a kapott hangzás unalmas, mesterséges lesz; nem imitálja egy valódi kórus gazdagságát. Ha azonban túlzásokba esünk, hasonló tapasztalatunk lehet, mint egy kevésbé tehetséges kórus hallgatása közben, amikor a szólamokon belüli eltérések miatt minden egyes előadót külön-külön hallhatunk (ezt hallhatjuk a 12.10. hangpéldában is).

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

12.2. Megvalósítás

A fázistolást a `HApp_12_01` (OSX) (WIN) példán keresztül vizsgáljuk (lásd a 12.4. ábrát). A késési időt (vagyis a fázistolás mértékét) frekvenciafüggő módon határozzuk meg, mégpedig úgy, hogy minden egyes frekvencia ugyanakkora arányú fázistolást szenvedjen el. Ezért nem időben, hanem fokokban fejezzük ki a fázistolás mértékét. Egy teljes periódus 360 fokból áll; ám mivel a szinuszhullámot fél periódussal (vagyis 180 fokkal) eltolva önmagának a -1-szeresét kapjuk, ezért a példánkban a fázistolás csak 0 és 180 fok között változtatható.



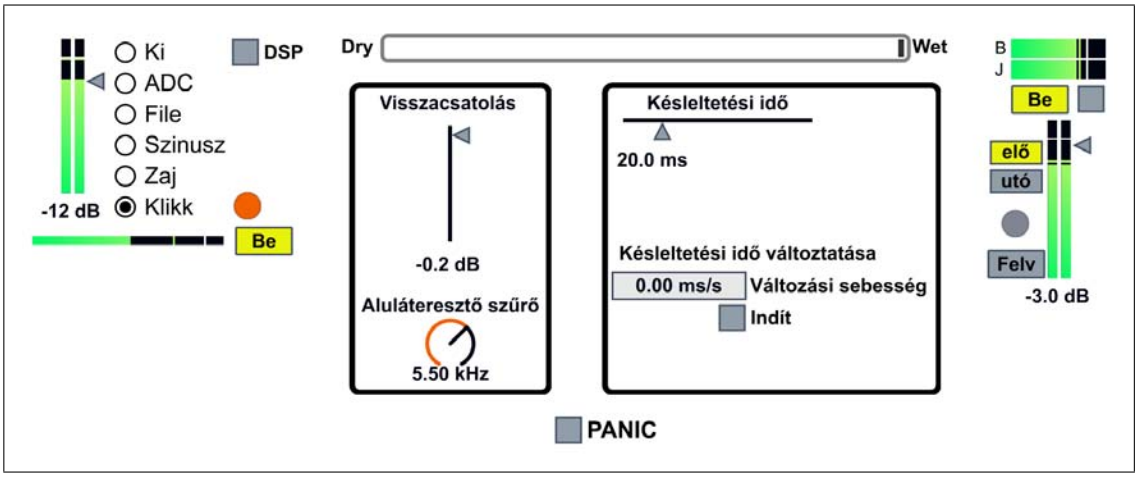
12.4. ábra. Fázistolás. A felső kijelző a bejövő, a középső kijelző a fázistolt jelet mutatja. A fázistolás mértékét fokokban adhatjuk meg. Az eredmény a legalsó kijelzőn jelenik meg.

Szinuszhullám esetén láthatjuk, hogy az eredmény egy, az eredetivel megegyező frekvenciájú szinuszhullám lesz, melynek amplitúdója annál kisebb, minél nagyobb a fázistolás. Ahogy fent említettük, a 180 fokkal eltolt szinuszhullám önmagának -1-szerese; emiatt 180 fokos fáziseltolásnál a fázistolt hullám kioltja az eredetit, így a kimenő jel elnémul. Mivel minden jel felbontható szinuszhullámokra, ezért tetszőleges jelet 180 fokkal eltolva (majd az eredetivel összeadva) csendet kapunk.

A `HApp_12_02` (OSX) (WIN), melyet a 12.5. ábra mutat, a késleltetés alapvető jellemzőit vizsgálja. A hatást a „*Dry/Wet*” csúszka segítségével keverhetjük az eredeti jelhez. A késleltetési idő a jobb oldali csúszkával szabályozható, míg a visszacsatolás erősítését a „*Visszacsatolás*” csúszka segítségével állíthatjuk be (ha tehát ez a hangerő $-\infty$ dB, akkor nincs visszacsatolás). A késleltetett jel egy „*Aluláteresztő szűrő*” közbeiktatása után csatolódik vissza a késleltető bemenetére. Amennyiben a rendszer begerjedne (ez akkor fordulhat elő, ha a késleltetett jel hangereje nemnegatív, és a visszacsatolás be van kapcsolva), a modul

Open Access 2026-11-02

alján található „*PANIC*” gombbal elnémíthatjuk a programot.



12.5. ábra. Késleltetés és visszacsatolás Max programnyelven.

A program a Doppler-effektust, vagyis a késleltetésen alapuló transzponálást is illusztrálja. A késleltetési idő folyamatos megváltoztatása a panel „*Késleltetési idő változtatása*” részén – a „*Változási sebesség*”[1] megadása után – az „*Indít*” feliratú gomb megnyomásával aktiválható. Mivel a modul kijelzőjén láthatjuk a késleltetési idő megváltozását, képet kaphatunk arról, hogy a késleltetésen alapuló transzponálás hogyan valósul meg.

A 12.1. videó szintén a Doppler-effektust illusztrálja, mégpedig egy mozgó sziréna hangjának szimulálása által. A mozgás során a sziréna a távolból „közelít felénk”, „elhalad előttünk”, majd újra „eltűnik a távolban”. A hatást egyetlen késleltető valósítja meg, melynek késleltetési idejét folyamatosan változtatjuk. A két kijelző egyrészt a pillanatnyi késleltetési időt, másrészt ennek változási sebességét (vagyis azt, hogy másodpercenként átlagosan hány ezredmásodperccel változik a késleltetési idő) mutatja. Láthatjuk, hogy a sziréna transzpozíciójának mértéke arányos ezzel a változási sebességgel: minél állandóbb e sebesség, annál kevesebbet változik a transzpozíció, illetve minél nagyobb a változás sebessége, annál nagyobb a transzpozíció mértéke. A sebesség előjele a transzpozíció irányát mutatja: negatív sebesség (csökkenő késleltetési idő) mellett a Doppler-effektus felfelé, míg pozitív sebesség (növekvő késleltetési idő) mellett lefelé transzponál.

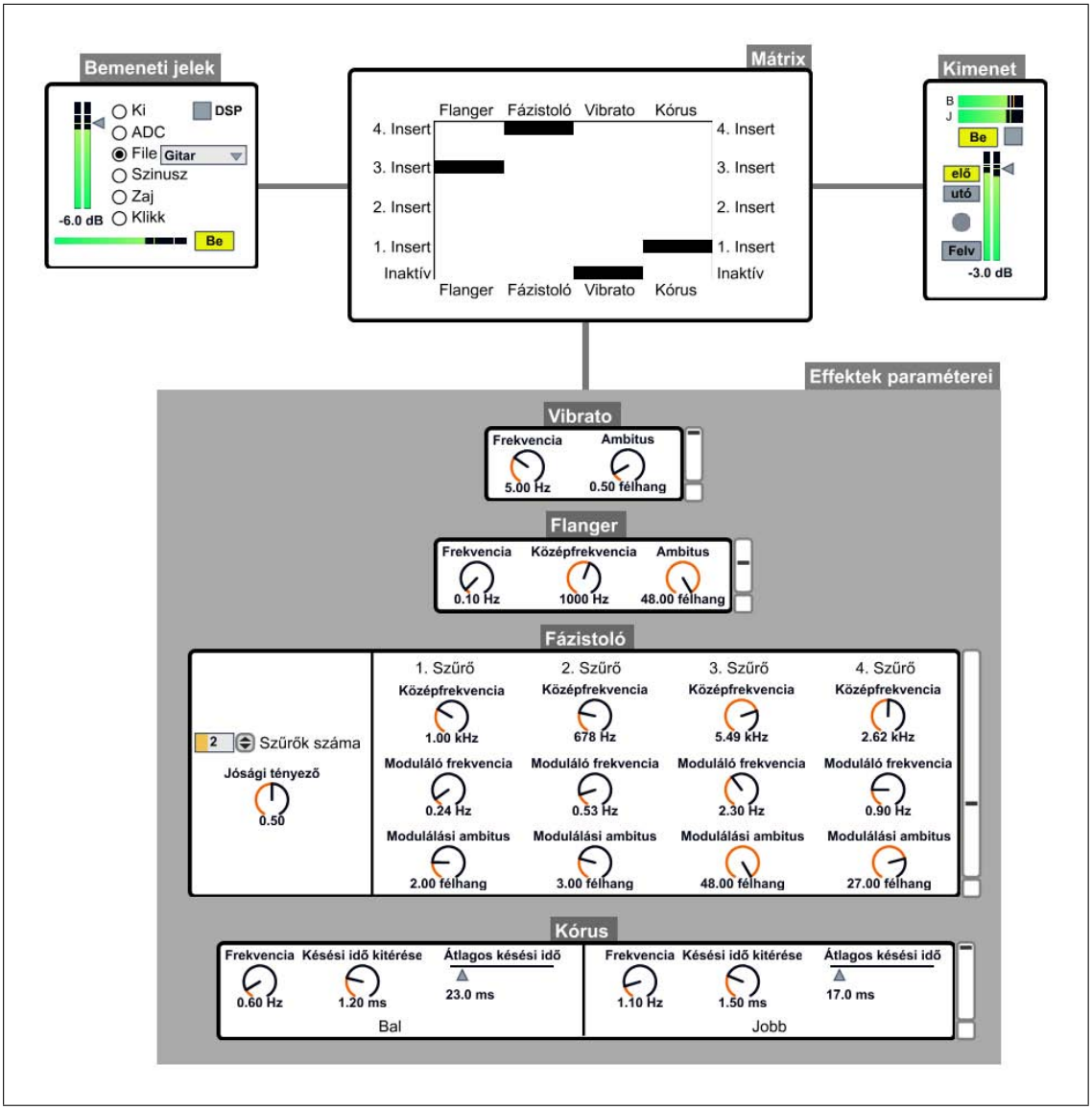
A fejezetben tárgyalt, szűrésre/késleltetésre épülő komplexebb modulokat a `HApp_12_03 (OSX) (WIN)` program (12.6. ábra) mutatja be. A beérkező hangot a felső sor közepén látható *mátrix* segítségével vezethetjük át az egyes jelfeldolgozó egységeken, mielőtt a kimenetre érkezne a hang. Az egyes effektek (Flanger, Fázistoló, Vibrato és Kórus) a mátrix oszlopaiban, míg az effektek sorrendje a mátrix soraiban helyezkednek el. Például a balról harmadik oszlop fentről második sorának aktiválásával a Vibrato effektet a 3. *insertbe* kötjük; ha az ezt megelőző insertekben nincs aktív effekt, akkor a hang először a Vibraton fog keresztülhaladni, majd ezt követően ér az utolsó insertbe (ha itt sincs aktív effekt, akkor a hang egyből távozik a kimeneten).

Minden egyes modul paramétereit egy-egy doboz tartalmazza. A dobozok jobb oldalán található vízszintes csúszkák az adott egység dry/wet szintjét állítják be, míg az csúszkák alatt látható kapcsolók aktiválása az adott modult teljesen kikapcsolja.

[1] Mértékegysége $\frac{\text{ms}}{\text{s}}$, mely az 1 másodperc alatti késleltetés-változást fejezi ki, ezredmásodpercekben.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



12.6. ábra. Késleltetésre/szűrésre épülő modulok Max programnyelven: Vibrato, Flanger, Fázistoló és Kórus.

A *Vibrato* modul két paraméterre épül, a vibrálás frekvenciájára és ambitusára. A modul hasonló elven működik, mint a 12.5. ábrán látott egyszerű késleltető, ám itt a késleltetési időt egy szinuszjel modulálja. Extrémül magas paramétereket választva a hatás az FM által elérhető hangzásokhoz hasonlít.

A *Flanger* modulban a korábban már leírt flanger-hatást próbálhatjuk ki. A „*Középfrekvencia*” a 12.3. ábra bal oldalán bemutatott spektrum első csúcsának helyét határozza meg; a „*Frekvencia*” és „*Ambitus*” beállítások a spektrum időbeli modulálását szabályozzák.

A *Fázistoló* egy (legfeljebb) négy, sorosan kapcsolt, mindent áteresztő szűrőt tartalmazó blokk. A szűrőket egy adott középfrekvencia jellemzi. A szűrők a spektrum középfrekvencia alatti részét változatlanul hagyják, míg a középfrekvencia feletti rész fázisát 180 fokkal eltolják. A jósági tényező itt azt határozza meg, hogy a 0 és 180 fok közötti átmenet a középfrekvencia körül milyen széles sávban történjen – minél nagyobb a jósági tényező, annál

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

élesebb az átmenet. Az egyes szűrők (külön-külön állítható) középfrekvenciáit külön-külön lehet egy-egy szinuszos jellel modulálni.

Amint említettük, nincs általános eljárás a kórus-effekt létrehozására. A panel alján található *Kórus* egy egyszerű sztereó implementáció, ahol a két csatorna (bal és jobb) fő paraméterei külön-külön állíthatóak. A program által megvalósított kórus-effekt külön-külön késlelteti a két sávot, mégpedig véletlenszerű mértékben. Ennek hatására a hang véletlenszerűen transzponálódik és szűrődik. A véletlenszerű késleltetés átlagos idejét, illetve az ettől való átlagos eltérést az „*Átlagos késési idő*”, illetve a „*Késési idő kitérése*” paraméterek szabályozzák. A „*Frekvencia*” paraméter azt határozza meg, hogy milyen gyakorisággal generálódjanak az új késési paraméterek a megadott időtartományon belül. Bizonyos átlagos késési idő fölött (nagyjából 40–50 ms) a kórushatás visszhangszerűvé válik; bizonyos (nagyjából 20 ms) idő alatt pedig a hatás egy véletlenszerűen változó szűrőhöz hasonlít. A késési idő kitérése elsősorban a kórus által alkalmazott vibrato mértékét befolyásolja.

12.3. Összegzés

Hangszín. A lineáris eszközök, habár új frekvenciákat nem hoznak létre, jelentősen képesek megváltoztatni a hangszínt. A megfelelő szűréssel fényessé vagy méllyé változtathatunk egy hangot, illetve rezonáns részeket hozhatunk létre a spektrumban. A késleltetés az egyszerű visszhangtól a bonyolult fázis-effektusokon keresztül ezernyi módon befolyásolja a hangzást. A paraméterek időben folyamatos változtatásával strukturálisan változtatható a teljes spektrum. Speciális (ám egyben a leggyakoribb) esetben, amikor a paramétereket valamilyen periodikus jellel moduláljuk, egy sor közismert zenei hatás érhető el. Ilyen a kórus, a vibrato, vagy éppen a hangszín-glissando (mely a flanger, illetve a fázistoló hatása).

Paraméterek. A legtöbb lineáris módszer kevés paramétert használ, és e paraméterek általában igen kézenfekvő módon jellemzik az egyes eszközöket. Ilyen például a késleltetési idő, vagy a szűrők karakterisztikáit leíró változók (pl. vágási frekvencia, jósági tényező). A paraméterek időbeli változtatásakor a modulálási ráta, illetve ennek erőssége (általában *mélységnek* hívjuk) zenei módon vezérli az adott hatást.

Műveletigény. Habár a különböző lineáris eljárások erőforrásigénye igen eltérő lehet, ez az igény általában elhanyagolható. Probléma legfeljebb azért merülhet fel, mert egy-egy komplex modul gyakran számos lineáris egységből épül fel, melyek műveletigénye ilyenkor összeadódik. Ám a lineáris komponenseket tartalmazó eszközök (pl. zengető) összesített erőforrásigénye általában nem jelentős.

12.4. Feladatok

1. Magyarázd meg (nagy vonalakban) a 12.3. ábra bal oldalán látható spektrum kialakulását a 12.1. folyamatábra alapján! Először idézd fel, hogy az azonos fázisban találkozó hullámok erősítik, az ellentétesben találkozók kioltják egymást! Második lépésként gondold meg, hogy ezek az erősítések és kioltások hogyan hozhatóak összefüggésbe az egyes frekvenciákhoz tartozó periódusidők, illetve a késleltetési idő viszonyával!

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

Más szóval: milyen relációban kell, hogy legyen a periódusidő a késleltetési idővel ahhoz, hogy az egyes frekvenciákon kioltás, illetve erősítés valósuljon meg?

2. A 6. fejezetben ismertettük a Karplus–Strong algoritmust. Implementáld a módszert a `HApp_12_02` `(OSX)` `(WIN)` program használatával! Milyen hangforrással tudnád a legjobban helyettesíteni az eredeti Karplus–Strong algoritmus által megkövetelt „véletlenszerű” kezdőállapotot? A választ a program különböző hangforrásaival történő kísérletezés útján add meg! Milyen paraméterekkel hozhatóak létre megpendített hangok? És dobszerűek? Mi az összefüggés a késleltetési idő és a kialakuló hang hangmagassága között?

© *Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola*

Open Access 2026-11-02

13. fejezet

Jelfeldolgozás II: torzítók

A torzítók a beérkező hang spektrumába új frekvenciákat illesztenek be. Kétféle módszert ismertetünk: az egyik a jel amplitúdójának átformálásán, a másik a mintavételezés megváltoztatásán alapul.

13.1. Elmélet

13.1.1. Jelalak-torzítás

A jelalak torzítása során a beérkező jel amplitúdóját skálázzuk át. Szinuszos jelek esetén (mivel az amplitúdók módosítása nincs hatással a periodicitásra) az újonnan beiktatott frekvenciák az eredeti frekvenciának egész számú többszörösei lesznek. Ezért az így létrehozható torzításokat *harmonikus torzításoknak* hívjuk. Komplex jelek esetén a felhangokon túl az eredeti jelben levő frekvenciák összegei és különbségei is megjelenhetnek (ez az *intermodulációs torzítás*).

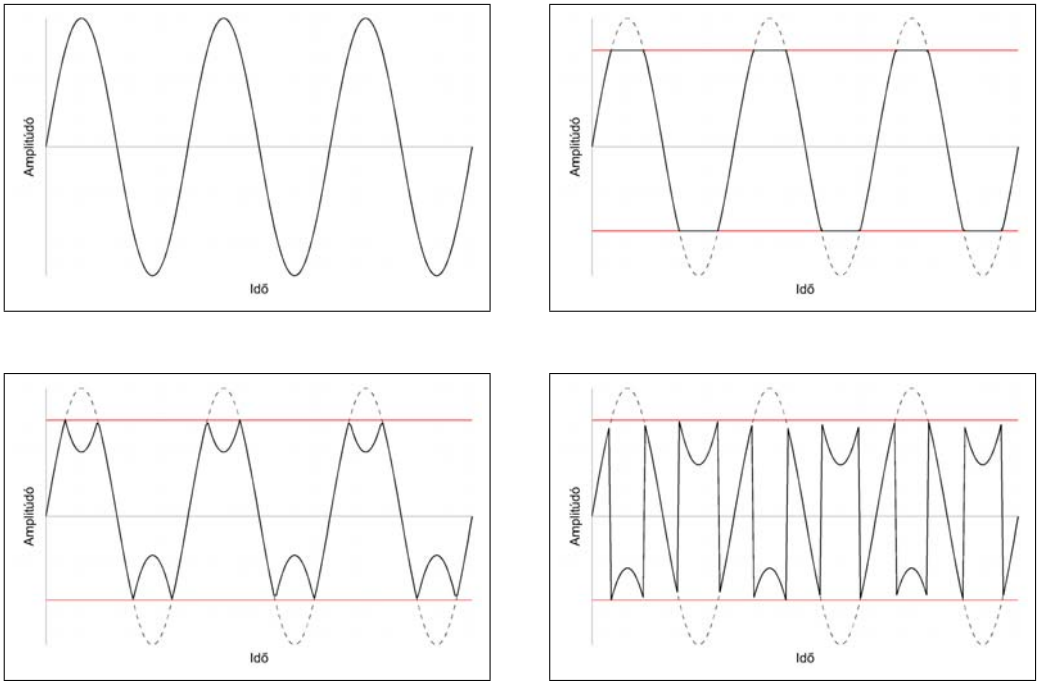
A legegyszerűbb módszer a *túlvezérlés*. Ekkor azokat a jelértékeket, melyek abszolút értéke meghalad egy bizonyos határt, „levágjuk” (a megadott határértékkel helyettesítjük). A levágás élességétől függően beszélhetünk *klippelésről* (akkor a határ alatt egyáltalán nem nyúlunk a jelhez, fölötte pedig teljesen levágjuk) és lágy torzításról. A *nullátmeneti torzítás* során a jel pozitív részéhez egy pozitív, negatív részéhez egy negatív konstanst adunk, így a jel két felét „széttoljuk” egymástól. *Körülfolyatáskor* a jelnek azt a részét, ami meghaladja a határértéket, „átmozgatjuk” a megengedett amplitúdó-tartomány aljára. *Tükrözéskor* a jelet „visszafordítjuk” a határról, a túlnyúló amplitúdórészt a határra tükrözzük. Ezek közül néhányat a 13.1. ábra illusztrál.

A fenti módszerek mindegyike egy általánosabb eljárás, a *hullámformálás* speciális esetei. Az eljárás során egy függvényt (*karakterisztikát*) kell megadni, ami minden lehetséges beérkező jelértékhez hozzárendel egy kimenő jelértéket. A torzítatlan kimeneti jelnek egy egyenes felel meg (ekkor ugyanis minden beérkező jelhez önmaga rendelődik); minél inkább eltér a karakterisztika az egyenestől, annál torzítottabbá válik a hang.

Igen speciális, ám nagyon hatásos eset, ha a karakterisztikát ún. *Csebisev-polinomokból* építjük fel. Egy egységnyi amplitúdójú szinuszhullámot egy n-edrendű Csebisev-polinom karakterisztikán átvezetve egy olyan szinuszhullámot kapunk, aminek frekvenciája az eredeti szinuszhullám n-szerese. Emiatt, ha a karakterisztikát különböző rendű Csebisev-polinomok összegéből képezzük, a végeredmény az eredeti frekvenciák felhangjait fogja kiadni. Ezért

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



13.1. ábra. Szinuszhullám (bal fent) különböző torzításai: klippelés (jobb fent), tükrözés (bal lent), körülfolyatás (jobb lent). A piros vonalak a levágási amplitúdókat jelölik.

az ilyen torzító kiválóan alkalmas az eredeti hangszín harmonikus telítésére. Az említett egységnyi amplitúdónak nagy szerep jut a hatás során. Ha a bejövő jel ennél hangosabb, az eredmény zajszerű torzítás lesz, míg nagyon halk bejövő jel esetén szinte semmi változást nem észlelhetünk.

Jelalak-torzítás hatására általában fényesebbé válik a hang, hiszen a szakadások és törések, amiket a hullámformába iktatunk, magas frekvenciák megjelenését okozzák; a technika nem alkalmas arra, hogy a hangzást sötétebbé tegyük, a mély regisztereket dúsítsuk. A jelalak-deformálás mindig növeli a jel zajosságát is, túlzott mértékben alkalmazva tetszőleges hangot zajjá változtat. A 13.2. ábra illusztrálja a fent bemutatott torzítási módszereket.

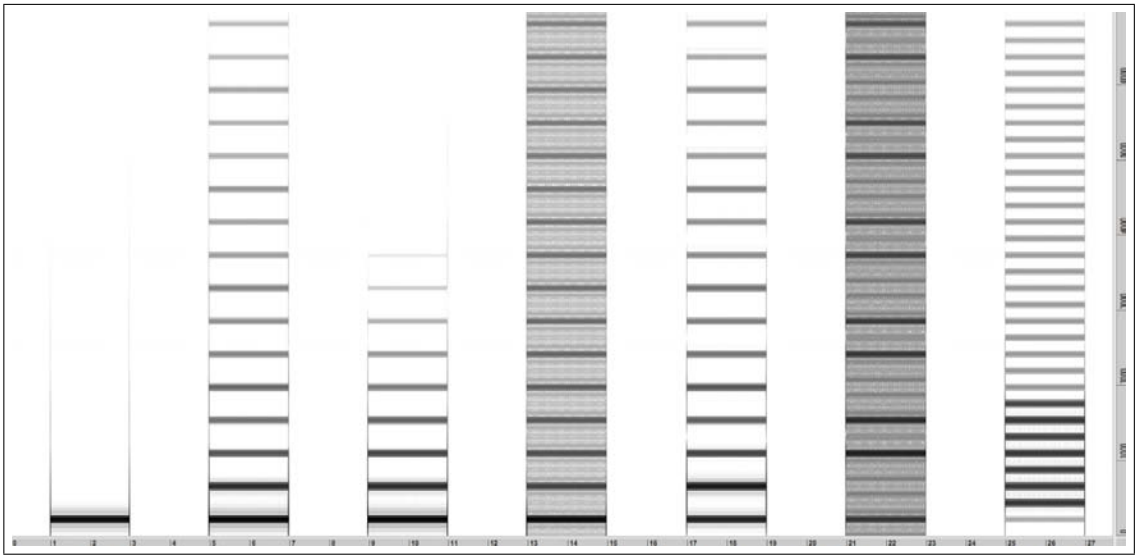
13.1.2. Kvantálási torzítás

A hang digitalizálása során azonos (nagyon rövid) időközönként (melynek neve *mintavételi idő*) mintát veszünk a hangból, és az így kapott értékek sorozatával reprezentáljuk az eredeti hangot. Az eljárás neve *frekvencia-kvantálás*, és a *mintavételi frekvenciával* (a mintavételi időnek, mint periódusidőnek megfelelő frekvencia) jellemezzük. Mivel a mintákat csak adott időközönként rögzítjük, ezért a digitális jelben csak olyan frekvenciák jelenhetnek meg, amelyek periódusideje hosszabb, mint a mintavételi idő. Pontosabban, a Nyquist–Shannon tétel értelmében egy digitális jel csak a mintavételi frekvencia felénél kisebb frekvenciákat tartalmazhat. Az ennél nagyobb frekvenciák álfrekvenciákat (lásd a 9.1.3. bekezdést) hoznak létre a digitális jelben.

Ráadásul a mintavételezett értékeket sem tudjuk végtelenül pontosan tárolni: a digitalizálás során a jelértékeknek számokat feleltetünk meg. A felbontás finomságát az így kapható legnagyobb és legkisebb szám aránya határozza meg. Az eljárás neve *dinamikai kvantálás*, melyet az imént említett aránnyal (pontosabban: az ennek megfelelő decibel-értékkel), a

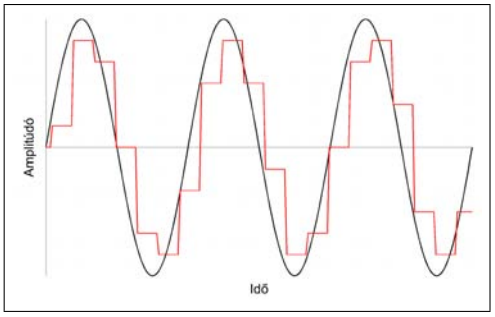
© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



13.2. ábra. Egy 220 Hz-es szinuszhullám és torzításai (szonogram). Balról jobbra: torzítatlan jel, klippelés, lágy torzítás, nullátmeneti torzítás, tükrözés, körülfolyatás, Csebisev-torzítás. A szonogram a 13.1. hangpélda alapján készült.

dinamika-tartománnyal jellemzünk. Mivel a jelben előforduló leghalkabb összetevő erejét a dinamika-tartomány nagysága határozza meg, ezért minél szűkebb ez a tartomány, annál hangosabbak lesznek a jelben kötelező jelleggel jelenlévő torzulások, zajok. Vagyis minél kisebb a dinamika-tartomány, annál zajosabb a jel. A frekvencia- és dinamikai kvantálást a 13.3. ábra illusztrálja.



13.3. ábra. Alacsony dinamikai felbontású szinuszjel. Az eredeti jelet a fekete, a kvantált jelet a piros vonal mutatja.

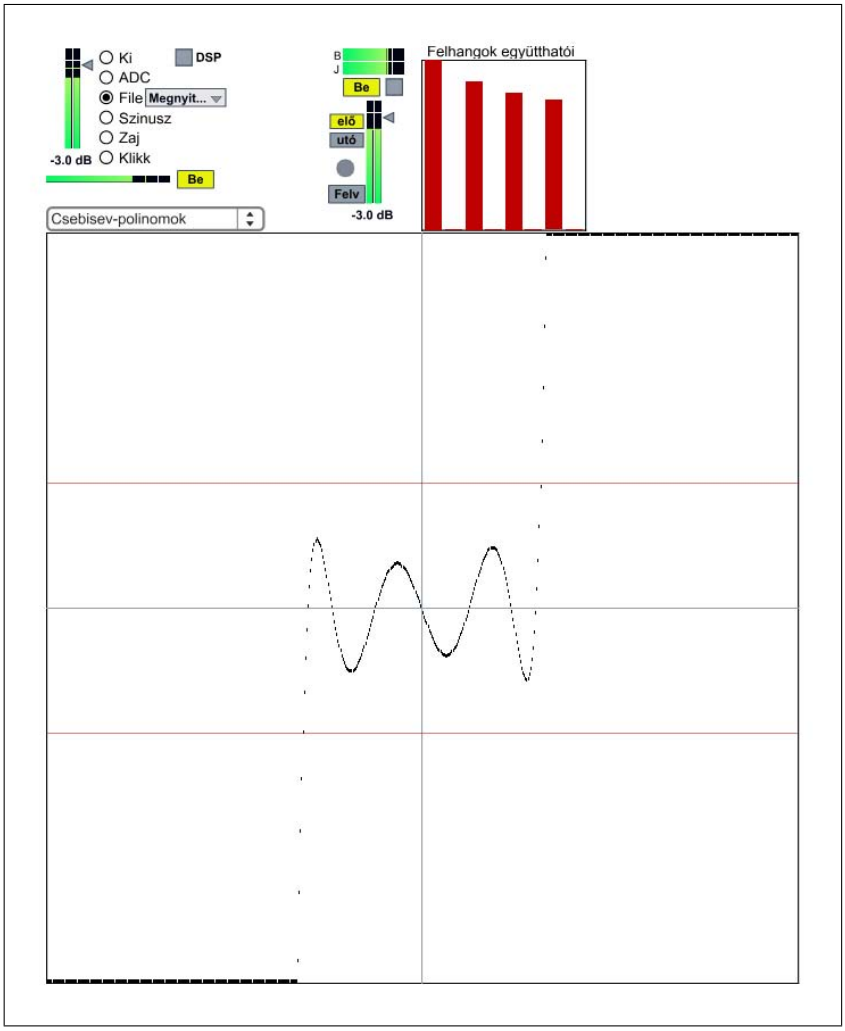
Ha egy korábban már digitalizált hangot „újra mintavételezünk”, a dinamika-tartomány szűkítésével zajossá, míg a mintavételi frekvencia csökkentésével – álfrekvenciák létrehozásával – inharmonikussá tehetjük. Ezt illusztrálja a 13.2. hangpélda, melyben egy zongoradallamot mintavételezünk újra. Megjegyezzük, hogy a módszer „visszafele” nem működik, egy zajos jelet nem lehet újra-mintavételezés útján „megtisztítani”.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

13.2. Megvalósítás

A jelalak-torzítást a `HApp_13_01` (OSX) (WIN) példán próbálhatjuk ki. A program kijelzője a torzítás karakterisztikáját mutatja. A karakterisztika bal felső sarkánál levő legördülő menüből választhatjuk ki a megfelelő torzítástípust; a kijelzőn ekkor az adott torzítást jellemző karakterisztika jelenik meg. Az „*Eredeti*” beállítás a jelet torzítatlanul engedi át, ezt a már említett egyenes karakterisztika jellemzi. A „*Lágy torzítás*”, illetve „*Nullátmeneti torzítás*” beállítások választásakor a menü mellett egy számdoboz jelenik meg, ahol a torzítás erőssége állítható. A „*Csebisev-polinomokat*” választva nyolc csúszka jelenik meg (ezt mutatja a 13.4. ábra), amivel az alaphang és az első hét felhang egymáshoz viszonyított amplitúdója jelenik meg. Amint a karakterisztikából látható, a harmonizátor-hatás akkor válik érzékelhetővé, ha a beérkező jel nem túl halk, de nem is túl hangos (vagyis amplitúdója 1 körül mozog). A „*Rajzolás*” opció választásával magunk szerkeszthetjük meg a karakterisztikát.

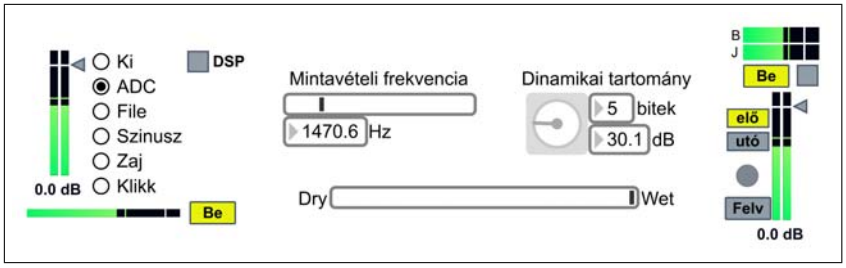


13.4. ábra. Jelalak-torzítás Max programnyelven. A központi kijelző a torzítás karakterisztikáját mutatja -3 és $+3$ között (lineáris amplitúdóértékek). A két vízszintes piros vonal a ± 1 amplitúdóértéket jelöli, a két szürke pedig az $x = 0$ és $y = 0$ tengelyeket.

A kvantálási torzítást a `HApp_13_02` (OSX) (WIN) program valósítja meg. A kezelőfe-

Open Access 2026-11-02

lület, melyet a 13.5. ábra mutat, nagyon egyszerű: a mintavételi frekvenciát (mely a hangkártya tényleges mintavételi frekvenciájánál nem lehet nagyobb) a bal oldali csúszkával és számdobozzal, míg a dinamikai tartományt a jobb oldali tárcsával, illetve számdobozzal szabályozhatjuk. Habár a dinamikai tartományt decibelben határozzuk meg, mivel a számítógép adott számú biten tárol egy-egy mintát, ezért tetszőleges decibel-érték nem adható meg. Így a példában csak az egy-egy mintára jutó bitek száma szabályozható.



13.5. ábra. Kvantálási torzítás Max programnyelven. A mintavételi frekvencia csökkentésével meg gyelhető az eredeti jel magas frekvenciáinak visszatükröződése, az álfrekvenciák megjelenése. A dinamikai tartomány szűkítésével a jel egyre zajosabbá válik.

Érdekes megfigyelni, hogy egy hosszan lecsengő jel dinamikai tartományának csökkentésével a lecsengés ideje (látszólag) megnő. Ez azért van, mert az eredeti lecsengés nagyobb dinamikai tartományban megy végbe, így idővel egyre kevésbé halljuk. A tartomány összenyomásával az a rész, amit az eredeti jelben már nem hallottunk, (rendszerint torzítva) újra hallhatóvá válik.

13.3. Összegzés

Hangszín. Mind a jelalak-, mind a kvantálási torzítás fényesebbé teszi a hangot. A jelalaktorzítás – körültekintő megvalósítást választva – dúsabbá, harmonikusabbá teheti a spektrumot, az újramintavételezés az álfrekvenciák megjelenése miatt inkább inharmonikus hangzások keltésére alkalmasabb. A túl erőteljes torzítás (függetlenül a választott módszertől) zajossá, recsegővé teszi a hangot.

Paraméterek. Habár a konkrét spektrum-deformációt csak a Csebisev-torzító esetén lehet viszonylag egyszerűen kiszámolni, a paraméterek hatása általában nagyon könnyen megérthető. Leggyakrabban egyetlen fő paraméterrel tudunk az eredeti és a fényesített, zajosított hangzás között interpolálni.

Műveletigény. Eljárástól függően lehet elhanyagolhatóan kicsi (pl. klippelés, tükrözés, körbefolyatás), de egyes eljárások nagyobb – habár nem extrém – számítási kapacitást igényelnek (pl. Csebisev-torzítás, újramintavételezés).

13.4. Feladatok

1. A Csebisev-polinomok – megfelelő amplitúdójú, szinuszos forrás használata esetén – ugyanolyan hangszínt hoznak létre, mint a 2. fejezetben bemutatott, szinuszok össze-

Open Access 2026-11-02

adásán alapuló hullámforma-ismétlési módszer. Határozzuk meg (kísérletezés segítségével), hogy milyen feltételek esetén tudjuk a hullámforma-ismétlést Csebisev-torzítással helyettesíteni! Mi történik, ha nem szinuszos forrást használunk? A beérkező hangnak milyen „optimális amplitúdó-tartományba” kell esnie, hogy a Csebisev-torzítás „karakteresen” – vagyis: ahol a torzítónak már jól érzékelhető hatása van, de még nem teszi (túlságosan) zajossá a beérkező hangot – szóljon?

2. Bizonyos beállítások mellett a nullátmeneti torzítás és a kvantálási torzítás hasonlóan (ám nem ugyanúgy) szól. Mi lehet ennek az oka? Hogyan kell a két torzítót beállítani, hogy a lehető leghasonlóbban szóljanak? Mi az a különbség, ami ilyenkor is megmarad a két torzító között? Milyen zenei helyzetben célszerű az egyik, és milyenben a másik megoldást használni?

3. Bizonyítsd be, hogy a kvantálási torzítást – amennyiben a mintavételezési frekvenciát nem változtatjuk meg – felfoghatjuk jelalak-torzításnak is: a `HApp_13_01` `(OSX)` `(WIN)` program „*Rajzolás*” opcióját választva rajzold meg egy 3 bites (-1 és 1 között összesen $2^3 = 8$ lehetséges amplitúdóértéket megengedő) kvantálási torzító karakterisztikáját! Hasonlítsd össze az eredményedet a `HApp_13_02` `(OSX)` `(WIN)` torzítójával ugyanazoknak a hangoknak a torzítása segítségével!

4. A `HApp_13_01` `(OSX)` `(WIN)` példa „*Klippelés (kemény torzítás)*”, „*Tükrözés*” és „*Körülfolyatás*” beállításainál a karakterisztika csak az 1-nél nagyobb abszolútértékű jelekre tér el a torzítatlan karakterisztikától, így ezek kipróbálásához olyan bemenő jelet kell választani, aminek az amplitúdója 1-nél nagyobb. Ezt legegyszerűbben a bemeneti erősítő megfelelő beállításával érhetjük el. A három torzítás tanulmányozása után hozz létre olyan klippelő, tükröző, illetve körülfolyató karakterisztikát a „*Rajzolás*” opció segítségével, amelyek már 1-nél kisebb amplitúdójú jelekre is hatnak!

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

14. fejezet

Jelfeldolgozás III: kompresszálás, zengetés

A fejezetben két, a keverés során alkalmazott alapvető technikát ismertetünk. A *kompresszálás* – avagy *(dinamikai) tömörítés* – során egy jel dinamika-tartományát csökkentjük. A *zengetés* célja egy adott hangtér szimulálása a térbeli visszaverődések, csillapítások és torzítások mesterséges előállítása segítségével.

14.1. Elmélet

14.1.1. Kompresszálás

A tömörítésnek (a felhasználás területétől függően) többféle célja is lehet. A telekommunikáció, illetve az adattárolási eljárások előszeretettel élnek a kompresszálás lehetőségével, hiszen a módszer csökkenti a jel, mint adatfolyam sávszélességét.[1] Hasonlóan fontos szerepet kap a tömörítés a stúdiókban, hiszen minél kisebb egy jel dinamika-tartománya, annál elenyészőbb a különbség a halk és a hangos tartományok között, ezáltal a halk részletek olyan környezetben is hallhatóak maradnak (pl. autórádió, szabadtéri hangosítás stb.), ahol tömörítés nélkül a háttérzaj elnyomná a halkabb hangokat. A kompresszióval emellett megváltoztatható egy-egy hangszer hangjának karaktere (pl. egy dobhang indulásának „élessége” csökkenthető, így – habár az ütések továbbra is jelen lesznek a végső hangzásban – az kevésbé fogja elnyomni a többi hangszert), amivel az egyszerre megszólaló hangszerek hangzását közelíthetjük egymáshoz.

A tömörítés egyenes következménye, hogy a hangzás „részletessége” csökken – erős kompresszálás esetén a hangzás finom részletei teljesen el is tűnhetnek. Alkalmazásakor ezért nagyon körültekintően kell eljárni.

A kompresszálás ellentéte az *expandálás*, mely során a dinamikai tartomány növelhető: a jel halk részei halkabbá és/vagy a hangos részei hangosabbá tehetők. Tipikus példa erre a *zajkapu*, amely egy bizonyos hangerő alatt elnémítja a jelet. A telekommunikációban gyakori, hogy a továbbítani kívánt jelet a továbbítás előtt kompresszálják, majd a célba juttatás

[1] Ha egy jel dinamika-tartományát felezzük, a kapott jel feleakkora tárolási kapacitást igényel az adathordozón, mint az eredeti. Ennek egyenes következménye, hogy egy adott adatátviteli eljárást választva a tömörített jel feleakkora átviteli sebességet igényel, mint az eredeti.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

után expandálják.[2] Az eljárás neve *kompandálás*.

A továbbiakban a tömörítést tárgyaljuk.

A kompresszálás alapjaiban véve egy igen egyszerű feladatból áll: amennyiben a jel túl halk, fel kell hangosítanunk, ha pedig túl hangos, le kell halkítanunk azt. A probléma bonyolultságának megértéséhez elég meggondolnunk, hogy ezt az egyszerűnek tűnő feladatot hogyan tudjuk megvalósítani? Néhány alapvető kérdés azok közül, amiket az előbbi eljárás felvet:

- Mekkora mértékben hangosítsuk, illetve halkítsuk a jelet? Vegyük figyelembe a beérkező jel pillanatnyi hangerejét, vagy (az eredeti hangerőtől függetlenül) a teljes jelet egy adott cél-hangerőre normáljuk?
- Mivel nincs zajmentes jel, a halk részek hangosításával a zajt is erősítjük. Hogyan döntsük el, hogy melyik halk részt érdemes erősíteni, és melyiket nem?
- Hogyan határozzuk meg a jel pillanatnyi hangerejét és milyen gyorsan reagáljunk a változásokra? Maga a jelfolyam csak azt mutatja meg, hogy az általa leírt rezgésnek mekkora a pillanatnyi kitérése az egyensúlyi helyétől (lásd az 1. fejezetet); a hangerőt a jel valamilyen átlagolásával számítjuk. A probléma abból adódik, hogy minél több mintát használunk ehhez az átlagoláshoz, annál kevésbé érzékeli a kompresszor a jel esetleges hirtelen ugrásait, ám ha túl kevés mintából határozzuk meg a hangerőt, a kompresszor túl érzékennyé válhat.[3]

Ahogy a fenti kérdésekből kitűnhet, a tömörítés többféle módon is megvalósítható. A legegyszerűbb esetben a beérkező jel hangerejét valamilyen egyszerű módszerrel megmérjük, magát a jelet pedig (amennyiben a hangerő túllép egy adott küszöböt) ez alapján halkítjuk. Természetesen olyan egyszerű konstrukció is elképzelhető, ami nem a megadott hangerőszint felett, hanem az alatt aktiválja a kompresszort – ebben az esetben nem halkítani, hanem erősíteni kell a jelet. A kétfajta megvalósítást a 14.1. ábrán mutatjuk be. A továbbiakban a két verzióról egyszerre beszélünk, úgy, hogy a folyó szövegben előforduló jelzők az adott hangerő felett levágó, míg a zárójelekben szereplők az adott hangerő alatt erősítő fajtára vonatkoznak.

A fent vázolt egyszerű, a jelet egy adott érték felett (alatt) gyengítő (erősítő) kompresszor-algoritmust a következő paraméterek jellemzik:

Tömörítési határérték: Ez az a referenciahangerő, ami felett (alatt) a kompresszor aktiválódik.

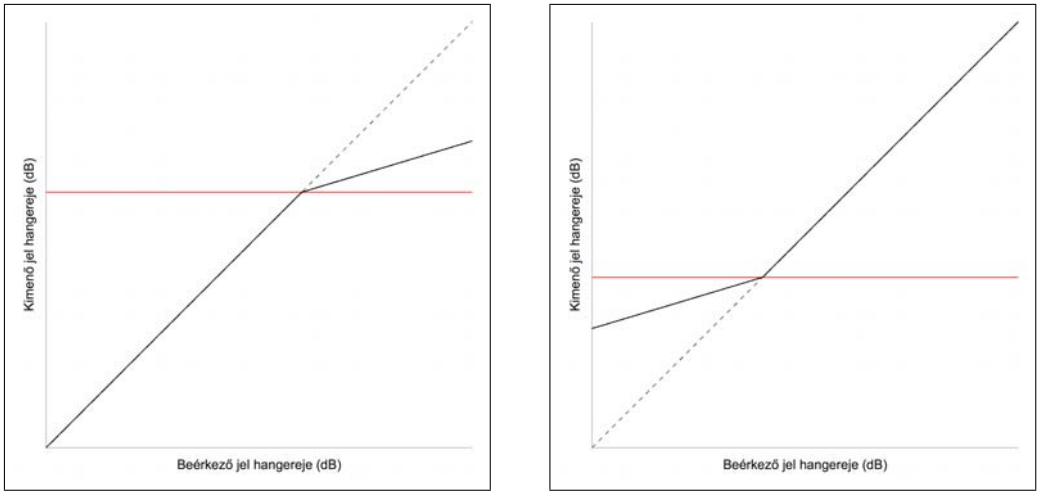
Tömörítési arány: Az eredeti jel hangerejének a referenciaszintet túllépő (alulmúló) részének aránya a kimeneti jel referenciaszintet túllépő (alulmúló) részéhez. Például, ha a tömörítési határérték -3 dB, a tömörítési arány 2:1 és a beérkező jel hangereje $+9$ dB (-15 dB), a kimenő jel hangereje $+3$ dB (-9 dB) lesz.

[2] Az így kapott végeredmény (amennyiben nem történik túlzott kompresszálás) hasonló lesz az eredeti jelhez, miközben a továbbításhoz felhasznált sávszélesség jóval kisebb is lehet, mint ha tömörítetlenül továbbítanák a jelet.

[3] Ha például egyetlen mintából határozzuk meg a hangerőt – vagyis a jel pillanatnyi amplitúdóértékét tekintjük hangerőnek – és „azonnal” reagálunk minden hangerőváltozásra, a kompresszor kimenetén egy konstans értékű jel fog megjelenni, hiszen a jel minden egyes mintáját ugyanarra az előre beállított amplitúdóértékre fogjuk normálni.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



14.1. ábra. Egyszerű kompresszorok sematikus karakterisztikái. A piros vonalak mindkét ábrán az aktiválódás határértékét, a feketék a végső hangerő-karakterisztikát mutatják. A bal oldali ábra a határérték felett halkít, a jobb oldali a határérték alatt erősít. Mindkét esetben meg gyelhető, hogy az eszköz egy nagyobb dinamikatartományt képez le egy kisebbre.

Felfutási idő: Ha a hangerő túllépi (alulmúlja) a referenciaszintet, a kompresszor elkezdi csökkenteni (növelni) annak hangerejét. A felfutási idő azt határozza meg, hogy[4] mennyi idő alatt éri el a kimenő jel a tömörítési arány által meghatározott új hangerőt; a paraméter gyakorlatilag a kompresszornak a referenciaszint átlépésekor tapasztalható „reakcióidejét” írja le.

Lecsengési idő: Ha a bejövő jel hangereje visszamegy a referenciaszint alá (fölé), a kompresszor kikapcsol. A hangerő folyamatossága miatt ez természetesen nem pillanatszerűen megy végbe. A lecsengési idő azt határozza meg, hogy[4] mennyi idő alatt ér vissza a kompresszor a nyugalmi helyzetébe.

A háttérzaj erősítésének elkerülése érdekében a legtöbb erősítő karakterisztikájú kompresszor nem kapcsol be, ha a beérkező jel hangereje egy adott (a tömörítési határértéknél lényegesen kisebb) határérték alá esik.

Ahogy utaltunk rá, a hangerő mérésének módja, illetve a kompresszor ezzel összefüggő „reakcióideje” nagyban befolyásolja az eszköz viselkedését. A mély hangok periódusideje sokkal nagyobb, mint a magasaké; ezért a spektrum felső regiszterében sokkal gyorsabban változhat a pillanatnyi hangerő, mint az alsó regiszterekben. Így egy egyszerűbb kompresszornál mindig kompromisszumot kell kötnünk: ha a felfutási, illetve lecsengési időt a mély hangok jellemző periódusidejéhez igazítjuk, azt kockáztatjuk, hogy a magas regiszterben lejátszódó folyamatokra az eszköz nem lesz képes megfelelően reagálni. Ezzel szemben, habár a rövid időállandók megfelelően követhetik a magas hangokat, eközben jó eséllyel „rángatni” fogják a mély regisztert, egyfajta véletlenszerű amplitúdó-modulációt létrehozva. Hogy illusztráljuk a problémát, gondoljuk meg, hogy míg egy pergődob *rimshotja* kevesebb, mint 100 ms-ig tart, egy nagybőgő vonóval történő megszólaltatásakor nagyjából ennyi idő szükséges pusztán ahhoz, hogy a stabil hang létrejöjjön. Az előbbi tömörítésekor tehát nagyon rövid, míg az utóbbi esetén nagyon hosszú felfutási időt célszerű választani.

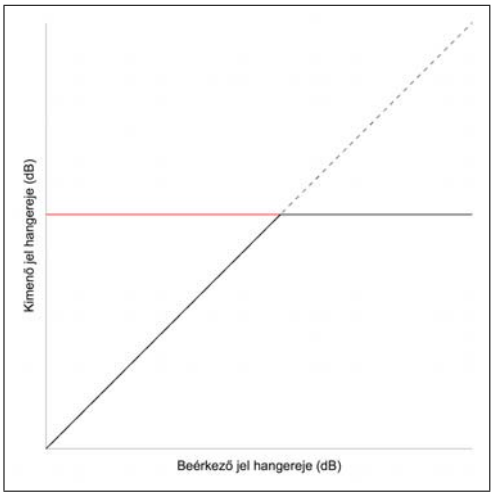
[4] Amennyiben az eredeti jel pillanatszerűen vált az új hangerőre, amit ezután konstans módon tart.

Open Access 2026-11-02

A legkézenfekvőbb megoldás a fenti problémára, ha a beérkező jelet egy szűrőrendszer segítségével több (független) komponensre bontjuk, és az így kapott sávkorlátolt jeleket külön-külön, komponensenként eltérő felfutási és lecsengési időkkel kompresszáljuk; a végeredményt a tömörített komponensek összeadásával képezzük.

A beszéd- és énekhangok tömörítésekor sokszor ez sem elég, ugyanis az egyes hangzók eltérő spektrális tulajdonságai miatt egy adott hangzót sokszor teljesen eltérő módon lenne célszerű kompresszálni, mint egy másikat. A legnagyobb problémát a zöngétlen mássalhangzók (c, cs, f, k, p, s, sz, t, ty) okozzák, ezek ugyanis nagyon rövid ideig tartanak, ám spektrumuk igen széles. Ezek megfelelő tömörítésére egy külön eszköz, a *de-esser* szolgál.

Az olyan kompresszort, melynek tömörítési aránya ∞:1, *limiternek* hívjuk (a ∞:1 karakterisztikát a 14.2. ábra mutatja be). A limiter tehát egy olyan eszköz, ami nem engedi, hogy a jel hangereje egy bizonyos határértéket átlépjen (alulmúljon). Ám a 13. fejezetben ismertetett klippeléssel szemben a limiter nem egyszerűen „levágja” a hullámforma határérték feletti (alatti) részét, hanem a teljes jelet lehalkítja (felerősíti). Ennek egyrészt az a következménye, hogy a hang sokkal kevésbé torzul, másrészt pedig az, hogy (amennyiben a beérkező hang) túl hirtelen ugrik, a limiternek nem lesz ideje időben reagálni; ez utóbbi miatt előfordulhat, hogy a limitált jel egyszer-egyszer átlépi a megadott határértéket, így a limitálás önmagában nem alkalmas a túlvezérlés teljes elkerülésére.



14.2. ábra. Adott határérték fölött levágó limiter karakterisztikája. A 14.1. ábra bal oldalán található karakterisztikával összehasonlítva látható, hogy a limiter valójában egy kompresszor, ahol a ∞:1 tömörítési arány miatt a karakterisztika határérték feletti része teljesen lapos.

Habár a limiter maga is egy kompresszor, gyakran előfordul, hogy a jelet (általános) kompresszálásnak és limitálásnak egyaránt alávetjük. Ennek több oka is van. Az egyik, hogy a limitálás, mint extrém tömörítési arányú kompresszálás, drasztikusabban befolyásolja a hangot, mint egy gyengébb kompresszálás, így zeneileg eltérő célokra használható (pl. limitálással sokkal könnyebb „elmosni” egy ritmikus zene hangindításait, mint gyenge kompresszálással). A másik ok a hangosítással kapcsolatos: mivel a limitált jel amplitúdója szinte sosem haladja meg a beállított határértéket, ezért limitálással biztosítható, hogy a hang

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

ne tegye tönkre az eszközeinket.[5]

A limitert – amennyiben a beérkező jelet egységesen (és konstans módon) felhangosítjuk, mielőtt átvezetnénk rajta – gyakran nem is a túl hangos részletek teljes elnyomása, hanem a halk részletek felerősítése miatt alkalmazzuk (ha ugyanis szimplán csak felerősítenénk a teljes jelet, úgy az eredetileg is hangos részletek túl hangossá válnának).

14.1.2. Zengetés

Képzeljünk el egy termet, amelyben egyetlen hangforrás és egyetlen hallgató van! Kézenfekvő feltételezés, hogy a hang azon a képzeletbeli egyenesen keresztül érkezik a hallgatóhoz, amely összeköti őt a forrással. Habár ez kétségtelenül így van, a hallgató nem csak ezt a „közvetlen" hangzást észleli. A hang ugyanis, miután visszaverődött a terem falairól, újra eléri a hallgatót. Sőt, ugyanez mondható el a visszavert hang visszaverődéseiről, és annak visszaverődéséről stb. is; e visszaverődések egészen addig tartanak, amíg a hang (egyrészt a visszaverődés, másrészt a terjedés során elszenvedett csillapítás miatt) előbb-utóbb teljesen el nem hal.

A trajektóriákat, amelyeket a hanghullámok a forrás és a hallgató között bejárnak, aszerint osztályozzuk, hogy hányszor verődnek vissza a terem falain (az osztályzást a 14.3. ábra illusztrálja). A „nulladrendű trajektória", a *közvetlen hang* egyenesen, visszaverődések nélkül jut a forrástól a hallgatóig. Az elsőrendűek egyszer, a másodrendűek kétszer verődnek vissza, és így tovább. A visszaverődések egyenként történő szimulálásával[6] elvileg jól közelíthetnénk, hogy egy adott térben elhelyezett forrást a tér egy másik pontjában levő hallgató pontosan hogyan fog észlelni.

Mivel egy ilyen számítás egyrészt túl bonyolult, másrészt függ a forrás és a hallgató konkrét helyétől (illetve a terem konkrét alakjától és egyéb fizikai tulajdonságaitól is), ezért a visszaverődéseket három csoportba soroljuk, s e csoportokat külön-külön kezeljük:

Közvetlen hang: a legtöbb zengetőben ez megegyezik a beérkező jellel (melyre esetleg egy rövid késleltetés kerülhet). Különleges szerepe annak tudható be, hogy általában lényegesen hamarabb érkezik a hallgatóhoz, mint bármelyik visszaverődés, így tudatunk elsősorban a közvetlen hang alapján határozza meg a hangforrás irányát (akkor is, ha ez sokszor halkabb, mint a később érkező visszavert hullámok sokasága).

Korai visszaverődések: A kifejezetten alacsony rendű visszaverődések tartoznak ide, melyek általában pár ezredmásodperccel követik a közvetlen hangot, és egyenként, még elkülöníthető módon érkeznek. Kulcsszerepet játszanak az adott hangtér méreteinek észlelésekor.

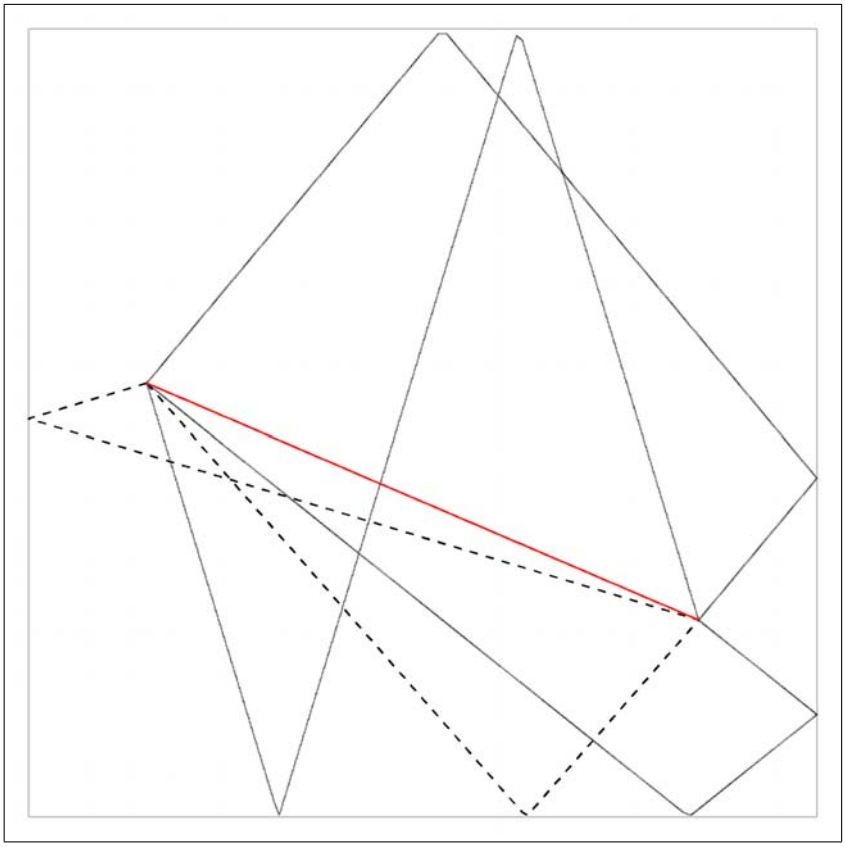
Utózengés: A sokszorosan visszavert hangokból épül fel. Ezek sztochasztikus módon keverednek, így nem alkalmasak a teremméret kifejezésére. A számos visszaverődés miatt az elszenvedett szűrési és torzítási hatások hatványozottan érvényesülnek e hangokban. Így ezek elsősorban a tér rezonanciáiról árulkodnak a tudatunk számára.

[5] Ahogy említettük, a túl hirtelen hangindításoknál a limitált jel egyszer-egyszer átlépheti a határértéket, ezért eszközeink tökéletes védéséhez a limitált jelet egy klipp-torzítón is át kell vezetni (lásd a 13. fejezetet), mielőtt a hangfalakra engedjük.

[6] Ez egyrészt késések beiktatásával (melyek a trajektória teljes bejárásához szükséges időt modellezik), másrészt szűréssel és torzítással (mely a terjedés, illetve a visszaverődés alatt elszenvedett csillapítást szimulálja) lehetséges.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



14.3. ábra. A hanghullámok által bejárható lehetséges trajektóriák egy része. A piros vonal a közvetlen hangot, a szaggatott feketék néhány elsőrendű, míg a pontozott feketék néhány másodrendű visszaverődést mutatnak.

A fenti csoportosítás önmagában csak a zengetés imitálásához szükséges hangok számát redukálja; továbbra is probléma, hogy ezeket a számításokat külön-külön el kell végeznünk minden egyes teremre, azon belül pedig minden egyes leheséges forráshelyre és hallgatóra. Ezért az általános zengető-modellek két alapfeltevésből indulnak ki:

- A korai visszaverődések nagyjából egyszerre érkeznek a hallgatóhoz, és így az egyetlen karakterisztikus idő, ami a teremméretet meghatározza, az a közvetlen jel és a korai visszaverődések között eltelő idő, melyet *előkésleltetésnek* hívunk. Mivel – amint már írtuk – a tudatunk a korai visszaverődések segítségével állapítja meg egy hangtér fizikai méretét, ezért számos zengető az előkésleltetés helyett egy virtuális „teremméret” paramétert kínál, melyből az előkésleltetést automatikusan számítja ki.

- Az utózengést kellően jól írja le, hogy a forrás kikapcsolása után mennyi ideig tart, amíg a visszaverődések megszűnnek. Ezt az időt – pontosabban azt az időt, ami alatt a visszaverődések átlagos hangereje az eredeti jel hangerejéhez mérten 60 dB-t esik – *lecsengési időnek* hívjuk. Szigorúan véve a lecsengési idő frekvenciafüggő, hiszen az utózengés jellegét nagyban befolyásolják a terem saját rezonanciái: a spektrum azon részei, amelyek e rezonanciákhoz közel esnek, sokkal lassabban csengenek le. Ezért a komolyabb zengetők lehetővé teszik, hogy (frekvenciasávonként) több, egymástól független lecsengési időt is megadhassunk.

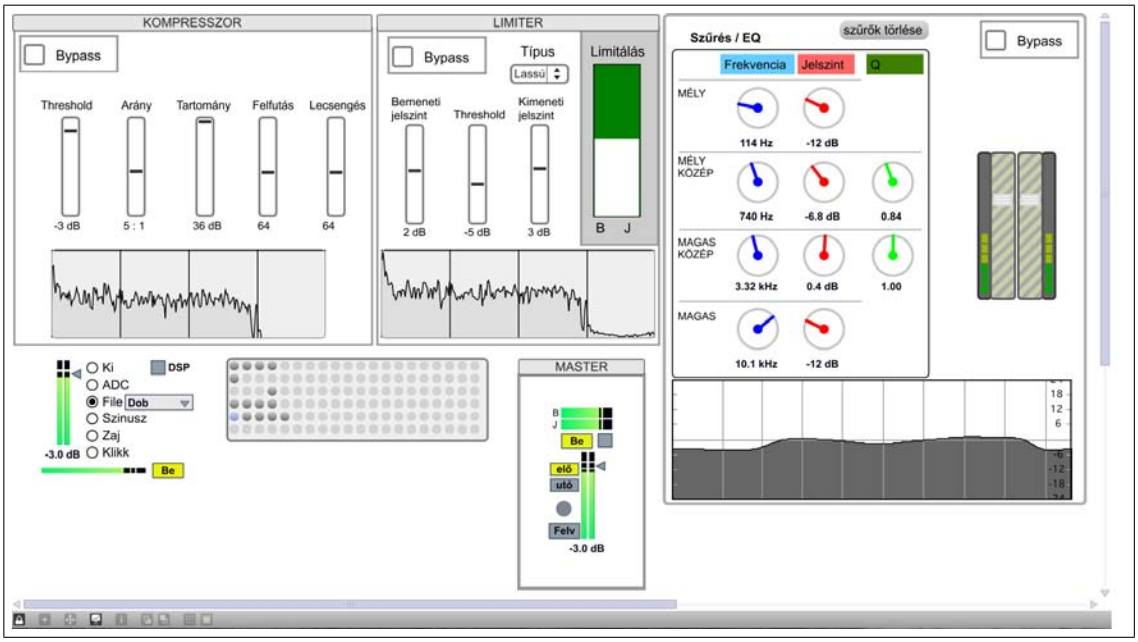
© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

Az egyszerűbb zengetők úgy modellezik az előkésleltetés, illetve a lecsengési idő frekvenciafüggését, hogy a zengetés során egy szűrőrendszer segítségével imitálják a terem rezonáns frekvenciáit. A legelterjedtebb zengetők ezért a (konstans jellegű) előkésleltetés és lecsengési idő megadása mellett néhány szűrőadat beállítását is lehetővé teszik.

14.2. Megvalósítás

A kompresszálást a `HApp_14_01` `(OSX)` `(WIN)` példaprogram illusztrálja (lásd a 14.4. ábrát). A bal felső panelen választhatunk forráshangot, mely először egy kompresszoron, majd egy limiteren, végül egy ekvalizátoron halad át.



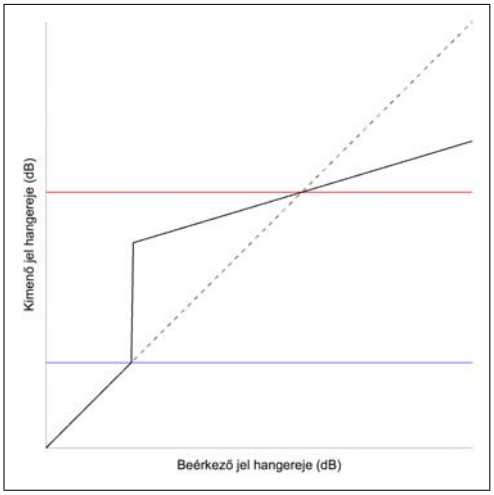
14.4. ábra. Kompresszálás, limitálás és ekvalizátor Max nyelven.

A kompresszor a fentebb említett két alapkonstrukciót egyesíti: amennyiben a hangerő túllépi a határértéket, halkítja, ám ha a határérték alatt marad a hangerő, erősíti azt. A határértéket a „*Threshold*”, a tömörítési arányt az „*Arány*”, míg a felfutási és lecsengési időket a „*Felfutás*” és „*Lecsengés*” szabályozókkal állíthatjuk. A háttérzaj erősítésének elkerülése érdekében beállítható egy minimális hangerőszint, amely alatt a kompresszor nem aktiválódhat. Ezt a „*Tartomány*” kapcsoló segítségével adhatjuk meg (az itt megadott hangerőérték a beállított küszöbértékhez képest mérendő). A kompresszor teljes karakterisztikáját a 14.5. ábra mutatja.

A limiter a levágási határértéken („*Threshold*”) felül két további beállítást is tartalmaz, melyek a bemeneti („*Bemeneti jelszint*”) és kimeneti („*Kimeneti jelszint*”) erősítőket szabályozzák. Ez lehetővé teszi a jel eredetileg halkabb szakaszainak az erősítését, mégpedig akkor, ha a bemeneti erősítő erősítése pozitív, a határérték pedig ennél alacsonyabbra van beállítva. A kimeneti erősítő a kapott jel normalizálását szolgálja. Az időállandókat a „*Gyors*” és „*Lassú*” beállítások határozzák meg. A limiter jobb oldalán látható kijelző mutatja, hogy mikor aktív a limiter, és az adott pillanatban mekkora mértékű levágást végez.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



14.5. ábra. A `HApp_14_01` `(OSX)` `(WIN)` programban alkalmazott tömörítő teljes karakterisztikája. A fekete vonal magát a karakterisztikát, a piros vonal a határértéket, a kék vonal pedig az aktiválódáshoz szükséges minimális hangerő szintjét mutatja.

Mind a kompresszor, mind a limiter egy-egy spektrogramon mutatja a kimenő jeleiknek a spektrumát. A 14.4. ábrán látható, hogy a limitálást követően a spektrum jóval laposabbá vált, mint eredetileg volt.

A kompresszor, a limiter, valamint az (a 4. fejezetben már bemutatott) ekvalizátor alkotják a *masterelés* alap-eszköztárát. Ezért a program jobb oldalán elhelyeztünk egy négycsatornás ekvalizátort is. A négy csatorna egy aluláteresztő, két sáváteresztő és egy felüláteresztő blokkból áll, mindegyik önálló frekvenciával és erősítéssel rendelkezik. Ezen felül a sávszűrők jósági tényezője is állítható. Az ekvalizátor teljes amplitúdó-karakterisztikáját a szűrő alján láthatjuk.

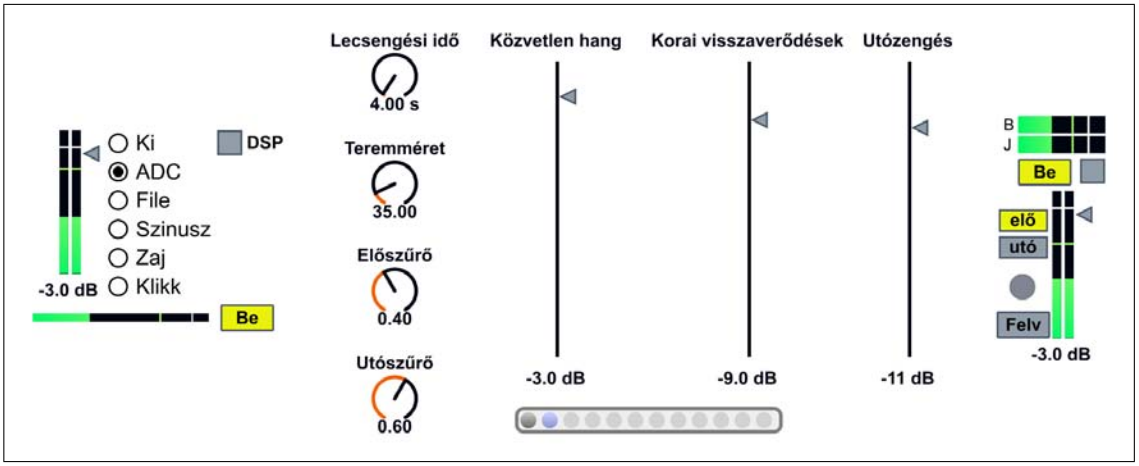
A program számos beállítást tartalmaz. Az első sorban lévő beállítások egyedül a kompresszort, a másodikban levő a limitert, a negyedikben levők pedig az ekvalizátort használják. A harmadik sorban egy példa illusztrálja a kompresszor és a limiter együttes alkalmazását, míg az utolsó sor beállításai mindhárom eljárásra támaszkodnak.

A `HApp_14_02` `(OSX)` `(WIN)` interaktív példa (lásd a 14.6. ábrát) egy egyszerű zengetőt tartalmaz. A korábban leírtak szerint külön-külön szabályozható a közvetlen hang, a korai visszhangok, illetve az utózengés. A szabályozást egyrészt a három különálló hangerőcsúszka teszi lehetővé, másrészt az előkésleltetés, illetve a lecsengési idő hangolhatósága. Amint említettük, a korai visszhangok szerepe elsősorban a terem méreteinek érzékelésekor játszanak szerepet; emiatt számos zengető (köztük a 14.6. ábrán látható is) az előkésleltetési idő konkrét megadása helyett a szimulálni kívánt terem méretét használja paraméter gyanánt, melyből a zengető maga határozza meg az optimális előkésleltetést. A zengető tartalmaz egy elő- és egy utószűrőt is, mellyel a visszaverődés (illetve a közegben való terjedés) során elveszített energia szimulálható.

A 14.1. CSound program egy fehérzajból és szűrőkből (vagyis szubtraktív szintézissel) előállított egyszerű hang zengetését és sztereó-pozicionálását mutatja be. A példa (melyet a 14.1. hangpélda tartalmaz) elején rövid, perkusszív hangok követik egymást, melyek térbeli elhelyezkedése és zengetése véletlenszerűen történik. A példa végén egy hosszú lecsengésű, zengetett zaj található.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02



14.6. ábra. Zengetés Max programnyelven.

14.3. Összegzés

Hangszín. A gyenge kompresszálás, illetve limitálás nem okoz akkora hangszínváltozást, mint a korábban ismertetett jelfeldolgozási eljárások. Habár a dinamika-tartomány csökkenése kiegyenlítettebbé teszi a hangerőviszonyokat, e folyamatok nem változtatják meg drasztikusan a spektrum struktúráját. Az erős kompresszió azonban többféle módon torzítja a hangszínt (a hangindítások élei elmosódhatnak, a háttér- és egyéb zajok kiemelődhetnek, egyes beállítások segítségével a hangok érdessé is válhatnak). A zengetésről hasonlót mondhatunk: ha a paramétereket valódi terek adataiból nyerjük, az elért hatás hasonló lesz, mint ha a beérkező jel, mint forráshang szólalna meg az adott térben. Ám a hangszín jelentős deformációja érhető el, ha extrém paraméterekkel (pl. irreálisan nagy lecsengési idő, teljesen lenémított közvetlen hang stb.) dolgozunk.

Paraméterek. A felépítés, illetve az egyes eljárások által alkalmazott modellek bonyolultságától függően a paraméterek száma és jelentése erősen eltérhet. Míg egyszerűbb kompresszorokat, zengetőket akár két-három változóval is jellemezhetünk, a bonyolultabb, többcsatornás eszközök paraméterigénye ennél jócskán nagyobb. Ráadásul legtöbbször a paraméterek számának növekedésével az egyes változók „zeneisége” csökken: míg egy kis paraméterigényű modell esetén a változók a „zenei összhatást” igyekeznek számszerűsíteni, a sokváltozós megoldások paraméterei közelebb állnak a „fizikai valósághoz” (pl. teremméret, felfutási idő stb.).

Műveletigény. Az egyszerűbb implementációk, melyek legfeljebb pár hangerőmérőt, szűrőt és késleltetőt tartalmaznak, nem igényelnek túl sok számítást; a bonyolultabb eszközök erőforrás-igénye igen nagy is lehet.

14.4. Feladatok

1. Vizsgáld meg a `HApp_14_01` `(OSX)` `(WIN)` program beállításainak első sorát, mégpedig legalább két különböző felvett hanggal! Ezek közül az egyik legyen perkusszív, a másik harmonikus jellegű (pl. dobszóló és kórus). Fogalmazd meg, hogy melyik be-

Open Access 2026-11-02

állítás hogyan befolyásolja a végeredményt! Az egyes beállítások különbségeit vesd össze a hallottakkal. Határozd meg ez alapján, hogy a végeredmény hogyan függ a kompresszor egyes paramétereitől!

2. Az előző feladathoz hasonlóan legalább két különböző hang segítségével elemezd a program beállításainak utolsó sorát! Minden egyes beállításnál vizsgáld meg, hogy mi változik, ha vagy a kompresszort, vagy a limitert kikapcsolod a *Bypass* kapcsolók segítségével! Ez alapján határozd meg, hogy a limiternek mi a szerepe az egyes beállításokban!

3. Hogyan befolyásolja a hangot az ekvalizátor? Különböző hangokkal és beállításokkal kísérletezve próbálj olyan eseteket találni, ahol az ekvalizátor hatása marginális, és olyanokat is, ahol az ekvalizátor nagymértékben befolyásolja a végeredményt! Lehetséges szabályokat állítani arra, hogy milyen esetekben juthat lényeges szerep az ekvalizátornak, és milyenekben nem?

4. Válassz ki egy tetszőleges hangfile forrást, majd állíts be számodra esztétikus kompresszor-, limiter- és ekvalizátor-értékeket! Nyiss meg egy olyan hangfile-t, ami lényegesen különbözik az előzőtől (pl. elsőre egy perkusszív, majd másodjára egy harmonikus hangmintát)! Továbbra is esztétikus a végeredmény? Miért?

5. Különböző beállításokon keresztül vizsgáld meg, hogy a `HApp_14_02` (OSX) (WIN) program elő- és utószűrője hogyan befolyásolja a zengetés hangszínét!

6. A kompresszor, mint zengető? A zengetett hang egyik tulajdonsága, hogy hosszabban cseng le, mint a nyers hang. Ezt a hatást kompresszorral is utánozhatjuk az időállandók és a határértékek extrém beállításaival. Kísérleti úton találj olyan paramétereket, amelyek a kompresszort egy „olcsó zengetővé” változtatják! Elérhető ugyanez a hatás a limiterrel is? Ha igen, mennyiben tér el a hatás a kompresszor által létrehozottól? Miben tér el egy zengetett hang hangszíne egy kompresszorral megnyújtott hangtól? Találj ki olyan zenei kontextust, ahol a kompresszor által keltett „olcsó zengetés” hatásosabb, mint az „igazi” zengetés!

14.5. Függelék

14.1. CSound-forráskód. `hd_14_01_csound.csd`

```
<CsoundSynthesizer>
<CsInstruments>

sr      = 44100
kr      = 4410
ksmps   = 10
nchnls  = 2

garvb       init        0

instr 1
; szűrt zaj + véletlenszerű panoráma és zengetés
        irvbsnd     linrand     .25
```

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02


```
        ipan        betarand    1, 1, 1
        kenv        linen       p4, p6, p3, p7
        kcut        expon       p8, p3, p9
        anoise      rand        p5
        afilt4      tone        anoise, kcut
        afilt3      tone        afilt4, kcut
        afilt2      tone        afilt3, kcut
        afilt1      tone        afilt2, kcut
        asig        =           afilt1*kenv
        outs        asig*ipan, asig*(1-ipan)
        print       ipan, irvbsnd
        garvb       =           garvb+(asig*irvbsnd)
endin

instr  199
        arvb        nreverb     garvb, p4, p5
        outs        arvb, arvb
        garvb       =           0
endin

</CsInstruments>
<CsScore>

        f 1 0 4096 10 1
        f 2 0 4096 10 1 .5 .333 .25 .2 .166 .142 .125 .111 .1 .09 .083
            .076 .071 .066 .062
        f 3 0 4097 20 2 1
        f 9 0 513 5 .001 128 .8 128 .6 256 1
        f 16 0 513 7 0 256 1 256 0
        f 17 0 513 7 0 10 1 246 1 10 0 246 0
        f 18 0 513 7 0 10 1 502 0

        ; p4=zengetési idő p5=magasfrekvencia szűrő
        i199 0 18 2.1 .8

        ; p4=amp p5=frq p6=atk p7=rel p8=cut1 p9=cut2
        i 1 0 .15 1  20000 .01 .01 10000 50
        i 1 +    .  <     .   .   .     <
        i 1 .    .  .     .   .   .     .
        i 1 .    .  .     .   .   .     .
        i 1 .    .  .     .   .   .     .
        i 1 .    .  .     .   .   .     .
        i 1 .    .  .     .   .   .     .
        i 1 .    .  .     .   .   .     .
        i 1 .    .  .     .   .   .     .
        i 1 .    .  .     .   .   .     .
        i 1 .    .  .     .   .   .     .
        i 1 .    .  .     .   .   .     .
        i 1 .    .  .     .   .   .     .
        i 1 .    .  .     .   .   .     .
        i 1 .    .  .5    .   .   .   100
        i 1 .    .  <     .   .   .     <
        i 1 .    .  .     .   .   .     .
        i 1 .    .  .     .   .   .     .
        i 1 .    .  .     .   .   .     .
        i 1 .    .  .     .   .   .     .
```

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02


```
69          i 1 .   .   .       .   .   .       .
70          i 1 .   .   .       .   .   .       .
71          i 1 .   .   .       .   .   .       .
72          i 1 .   .   .       .   .   .       .
73          i 1 .   .   .       .   .   .       .
74          i 1 .   .   .       .   .   .       .
75          i 1 .   .   .       .   .   .       .
76          i 1 .   .   .       .   .   .       .
77          i 1 .   .   .       .   .   .       .
78          i 1 .   .   .       .   .   .       .
79          i 1 .   .   .       .   .   .       .
80          i 1 .   .   .       .   .   .       .
81          i 1 .   .   .       .   .   .       .
82          i 1 .   .   .       .   .   .       .
83          i 1 . 3   1         .   .   . 10000
84
85  </CsScore>
86  </CsoundSynthesizer>
```

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola

Open Access 2026-11-02

15. fejezet

Multidimenzionális hangszínvektorok

Az eddigiekben ismertettük a leggyakoribb hangszintézis- és modulációs technikákat (2–9. fejezetek), illetve a legfontosabb hangátalakítási eljárásokat (11–14. fejezetek). Láthattuk, hogy melyik módszerrel milyen hangszíneket állíthatunk elő, illetve az egyes eljárások lehetséges paraméterei hogyan képesek befolyásolni az előállítandó hangszínt.

A zenei logika azonban fordítva működik: a kiindulási pont általában egy-egy adott hangszín, illetve hangszíntérszegmens, a cél pedig a megfelelő eszközök kiválasztása, az egyes módszerek közti viszonyok (paraméter-kapcsolatok) optimális meghatározása.

Természetesen a megfelelő hangszínek és hangszínkapcsolatok megtalálása egy olyan zeneszerzői feladat, amit egyedileg, az adott zenemű sajátosságainak figyelembevételével kell megoldani, és mint ilyen, messze túlmutat ennek a tananyagnak a keretein. Léteznek azonban stratégiák, amelyek mentén elindulhatunk.

A korábbiak során láthattuk, hogy egyazon hangszínhez több úton is eljuthatunk. Zenei szempontból a különbség az egyes módszerek között abban mutatkozik meg, hogy a hangszíntér adott pontjából milyen irányokban, mekkora mértékben és milyen nehézségek árán mozdulhatunk el. Az egyes szintézis- és jelfeldolgozási módszerek kiválasztásakor tehát mindkét szempontra (az elérendő hangszínekre és a köztük vezető utakra egyaránt) tekintettel kell lennünk.

15.1. Elmélet

15.1.1. Hangszínvektorok

A hang magasságával, vagy időtartamával ellentétben a hangszínt nem lehet egyetlen paraméterrel jellemezni. Ahogy az 1. fejezetben utaltunk már rá, a hangszínteret a hangszíntér-dimenziók sokaságával írjuk le. Amikor tehát egy adott hangszínt meg kell határoznunk, ideális esetben az összes releváns hangszíntér-dimenzió megadásával tesszük azt.

Az egyes hangszíntér-dimenziók értékeit egy vektorban, a *hangszínvektorban* soroljuk fel. A hangszínvektor nem más, mint a multidimenzionális hangszíntér egy adott pontja; a (folytonos) hangszínváltozásokat úgy foghatjuk fel, mint egy útvonalat a multidimenzionális hangszíntérben, melynek minden egyes pontjához különböző hangszínvektorokat rendelhetünk. A kompozíciós folyamat (legalábbis a felhasznált hangszínek szempontjából) nem más, mint azon pontok és útvonalak megadása, amelyeket a darab során be kell járnunk.

© Szigetvári Andrea, Siska Ádám, Budapesti Kommunikációs és Üzleti Főiskola