

Wavelet transzformáció

Számítógépes szimulációk, 1. projekt

Asztalos Bogdán

2019. október 25.

Tartalomjegyzék

1. Bevezető	2
2. A wavelet transzformáció elméleti háttere	2
3. Ismerkedés a wavelet transzformáció használatával	4
3.1. Példajel elemzése	4
3.2. Folytonosan változó frekvencia	6
3.3. Tanulságok az ismerkedés alapján	6
4. D-dúr kánon szolmizálása	8
5. Dallam készítése spektrum alapján	10
6. Összegzés	12

1. Bevezető

A hagyományos Fourier-transzformációval felbonthatunk egy adott jelet harmonikus komponensekre, azaz felírhatjuk a jelet $\cos(\omega t)$ és $\sin(\omega t)$ alakú függvények összegeként. Ez azonban csak véges tartományon, vagy periodikus jel esetében tehető meg, és a transzformációval kapott együtthatók az egész tartományra érvényesek. Ez egy rendkívül hasznos, és elméleti jelentőséggel is bíró matematikai eszköz, ám ha a vizsgálandó jelünk nem stacionárius, akkor nehezen alkalmazható.

A wavelet transzformáció a Fourier-transzformáció általánosításaként értelmezhető oly módon, hogy megengedi a jelben előforduló frekvenciamódusok időfüggését. Ezt úgy tudjuk megtenni, hogy a transzformáció során a bázisfüggvénynek nem egy teljes tartományon oszcilláló szinusz vagy koszinusz függvényt használunk, hanem egy hullámcsomagot (angolul *waveletet*), amely csak lokálisan rezeg, és rövid idő alatt lecseng, ezáltal a jel és a hullámcsomag konvolúciója az adott módus időfüggő jelenlétét fogja jellemezni.

A wavelet transzformációt, a Fourier-transzformációhoz hasonlóan, számos helyen alkalmazzák az asztrofizikától az elektronikáig; elsődleges funkciója, hogy változó frekvenciájú rezgések által keltett jeleket dolgozzanak fel a segítségével. Ugyanakkor, nem csak a tudományos világban kell keresni időben változó frekvenciájú jeleket, hiszen az emberi hallás is azon alapszik, hogy a levegő adott frekvenciájú rezgéseit észleljük fülünkkel. Ahhoz, hogy az emberi beszédet megértsük, vagy a hallgatott zenét élvezzük, agyunknak folyamatosan fel kell dolgoznia azt, hogy hogyan változik az érzékelt hanghullámok frekvenciája, vagyis ha agyunk hallásközpontját akarjuk modellezni, akkor is a wavelet transzformációt kell használnunk.

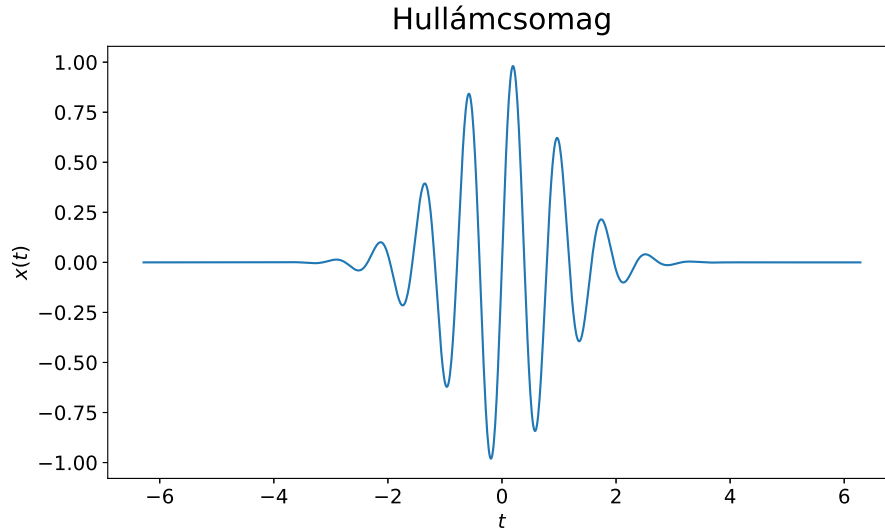
Az alábbi munkában a wavelet transzformáció alapvető működésével ismerkedünk meg egy igen egyszerű algoritmus segítségével, majd egy hanganyagot elemzünk a használatával, végül pedig megpróbálunk egy saját hangmintát létrehozni az inverz wavelet transzformációval.

2. A wavelet transzformáció elméleti háttere

A wavelet transzformáció olyan bázissal dolgozik, amely hullámcsomagokat tartalmaz. Ezeket a hullámcsomagokat általában egy Ψ alapfüggvénnyel lehet definiálni, amely a 0 körül oszcillál és nagy távolságokban gyorsan lecseng. Precízen, az alapfüggvénynek az alábbi feltételeket kell kielégítenie:

1. $\Psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ vagyis Ψ legyen abszolút- és négyzetesen integrálható:

$$\int_{-\infty}^{\infty} |\Psi(t)| dt < \infty \quad \text{és} \quad \int_{-\infty}^{\infty} |\Psi(t)|^2 dt < \infty \quad (1)$$



1. ábra. A $\Psi(t) = \sin(8t)e^{-t^2/2}$ képlettel megadott alapfüggvény, amely megfelel a szükséges feltételeknek.

2. Legyen 0 átlagú:

$$\int_{-\infty}^{\infty} \Psi(t) dt = 0 \quad (2)$$

3. Legyen 1-re normált:

$$\int_{-\infty}^{\infty} |\Psi(t)|^2 dt = 1 \quad (3)$$

4. Általában technikai szempontból hasznos, ha az alacsony m -ik momentuma 0:

$$\int_{-\infty}^{\infty} t^m \Psi(t) dt = 0 \quad (4)$$

Egy jó választás például az alábbi hullámcsomag (amit a későbbiekben is használunk majd), ahol a közönséges szinusz hullámot a Gauss-függvénnyel súlyozzuk:

$$\Psi(t) = \sin(8t)e^{-t^2/2} \quad (5)$$

Adott Ψ alapfüggvénnyel a wavelet transzformáció az alábbi módon értelmezendő: $y(t)$ jel transzformáltja:

$$Y(s, \tau) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} \Psi^* \left(\frac{t - \tau}{s} \right) y(t) dt \quad (6)$$

Az inverz transzformáció:

$$y(t) = C \int_{s=-\infty}^{\infty} \int_{\tau=-\infty}^{\infty} \frac{1}{\sqrt{|s|}} \Psi^* \left(\frac{t - \tau}{s} \right) Y(s, \tau) d\tau \frac{ds}{s^2} \quad (7)$$

ahol C egy konstans érték. A transzformált $Y(s, \tau)$ függvény azt adja meg, hogy a jel $t = \tau$ időpillanatában mennyire jellemző rá az $\omega = \frac{2\pi}{s}$ frekvenciamódus. s jelentése tehát egy effektív periódusidő: minél nagyobb az értéke, az oszcillálás annál lassabb. [1, 2]

3. Ismerkedés a wavelet transzformáció használatával

Az alábbi munka alapjául R. H. Landau könyvének 11. fejezete szolgált [1], amelyben a wavelet transzformáció alapjainak a megismerése mellett hoz egy példaprogramot a wavelet transzformáció implementálására. A könyvben lévő 11.1 kód egész pontosan azt a feladatot oldja meg, amelyet a 3.1 alfejezetben vizsgálunk majd, ezért ahhoz, hogy tetszőleges jelet feldolgozzunk vele, az algoritmusokat kissé általánosabbá kellett tenni. A wavelet transzformációt és az inverz wavelet transzformációt végző kódrészleteket függvényekbe rendeztem, amelyek a jelen munka mellé csatolt munkafüzetek elején, a Függvények részénél megtalálhatók.

Ebben a fejezetben két példán keresztül megismerkedünk ezen algoritmusok használatával, hogy a későbbiekben tényleges alkalmazásokhoz is használhassuk őket.

3.1. Példajel elemzése

A könyvben az alábbi jelet használják példaként, amin keresztül bemutatják a wavelet transzformáció alkalmazását:

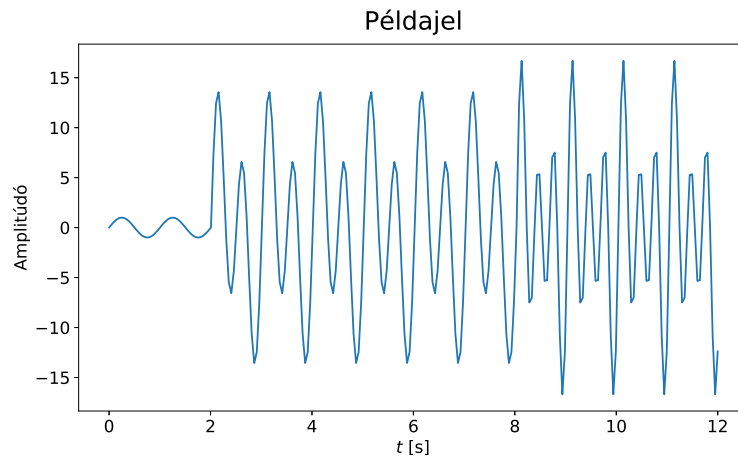
$$y(t) = \begin{cases} \sin 2\pi t & \text{ha } 0 \leq t \leq 2 \\ 5 \sin 2\pi t + 10 \sin 4\pi t & \text{ha } 2 \leq t \leq 8 \\ 2, 5 \sin 2\pi t + 6 \sin 4\pi t + 10 \sin 6\pi t & \text{ha } 8 \leq t \leq 12 \end{cases} \quad (8)$$

Elsőként azt ellenőriztük, hogy visszakapjuk-e a könyvben látott eredményeket a wavelet transzformációt használva.

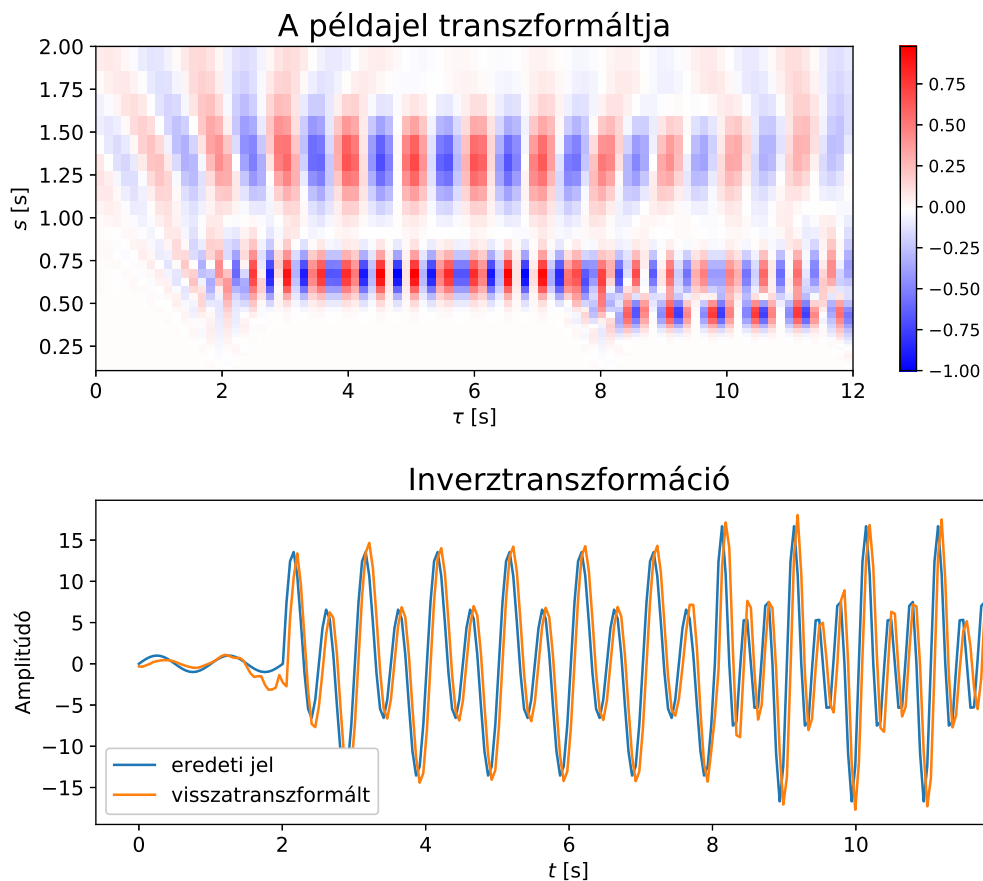
A jel elemzésének kódja a mellékletben csatolt `notebooks/warm-up.ipynb` fájl első felében található. Két dolgot csináltunk: Először alkalmaztuk a (6) kifejezésben definiált transzformációt az $y(t)$ jelre, ebből megkaptuk a $Y(s, \tau)$ transzformált jelet. Ezután pedig erre a transzformált jelre alkalmaztuk a (7) inverz transzformációt, amiből az $\tilde{y}(t)$ inverz transzformált jelet kaptuk meg. A transzformáció és az inverz transzformáció eredményét a 3. ábrán ábráztuk.

A transzformált jel ábráján jól megfigyelhető a három vízszintes nyom, amelyek a rezgés három módusához tartozik. Mivel a különböző módusok különböző időpillanatokban jelennek meg, ezek a vízszintes elszíneződések is különböző hosszúságúak.

A wavelet transzformáció tulajdonságai alapján, a visszatranszformált jelnek meg kell egyeznie az eredeti jellel. Ahogy a 3. ábra alsó felén látható, ez nem egészen történt



2. ábra. A (8) egyenlettel definiált példajel alakja. Jól látható, hogy a függvény három szakaszra bontható, ahol a módusok különböző módon jelennek meg.



3. ábra. *fent*: A példajel wavelet transzformáltja. Jól elkülöníthető a rezgés három szakasza, és az, ahogy a három módus különböző időpillanatokban jelennek meg. *lent*: A példajel, és a transzformáltjának inverz transzformáltja. A két jel majdnem pontosan fedik egymást.

meg. Ennek oka az, hogy az inverz transzformáció csak folytonos transzformáció esetén, azaz tényleges integrálás esetén lesz valóban az inverz. Mivel itt az integrálás helyett diszkrét összegzést végeztünk, az eredmény nem pontos, ám szemre megállapítható, hogy visszatranszformált minta majdnem fedí az eredetit.

3.2. Folytonosan változó frekvencia

Annak érdekében, hogy ne csak az előbb vizsgált tankönyvi példán keresztül ismerkedjünk meg a transzformációval, megvizsgáltunk egy másik, szintén elég alapvető problémát, amely a `notebookswarm-up.ipynb` fájl második felében található. A vizsgált jel a

$$y(t) = \sin(0.15t^2) \quad (9)$$

volt. Ez értelmezhető úgy is, mint egy változó $\omega(t) = 0.15t$ frekvenciájú rezgés. Eszerint, ha erre alkalmazzuk a wavelet transzformációt, akkor a legerősebb módus τ függvényében az $s(\tau) \sim \omega^{-1}(\tau) \sim \tau^{-1}$ szerint fog változni.

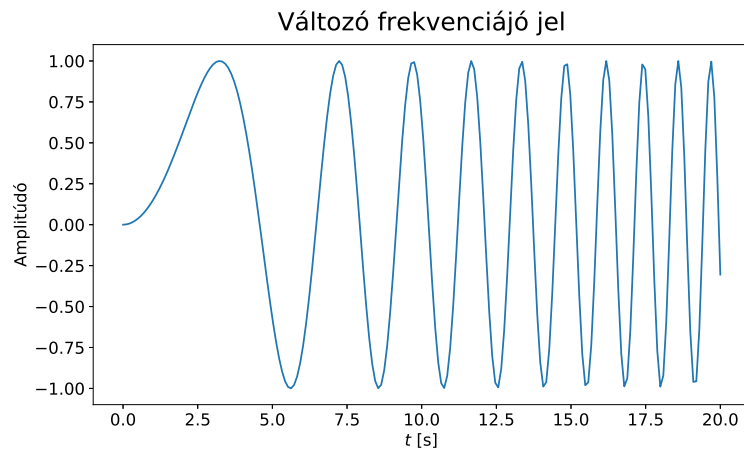
Az elvégzett munka ebben az esetben is az $y(t)$ jel transzformálása és visszatranszformálása volt. Ennek eredménye az 5. ábrán látható.

A transzformált jel $s - \tau$ síkján megjelöltük egy zöld görbével a $s \sim \tau^{-1}$ kapcsolatot. Mivel a (9) képlet alapján ilyen az elmélet összefüggés s és τ között, ezért azt várnánk, hogy a mérvadó módusok ezen görbe közelében helyezkednek el. Ez valamilyen szinten igaz is, hiszen a görbe vonulata nagyjából megfigyelhető a transzformált jelben is, de a spektrum nagy s -ek esetén jelentősen szétkenődött, így nem jelenik meg annyira pontosan az effektus, amire számíthattunk.

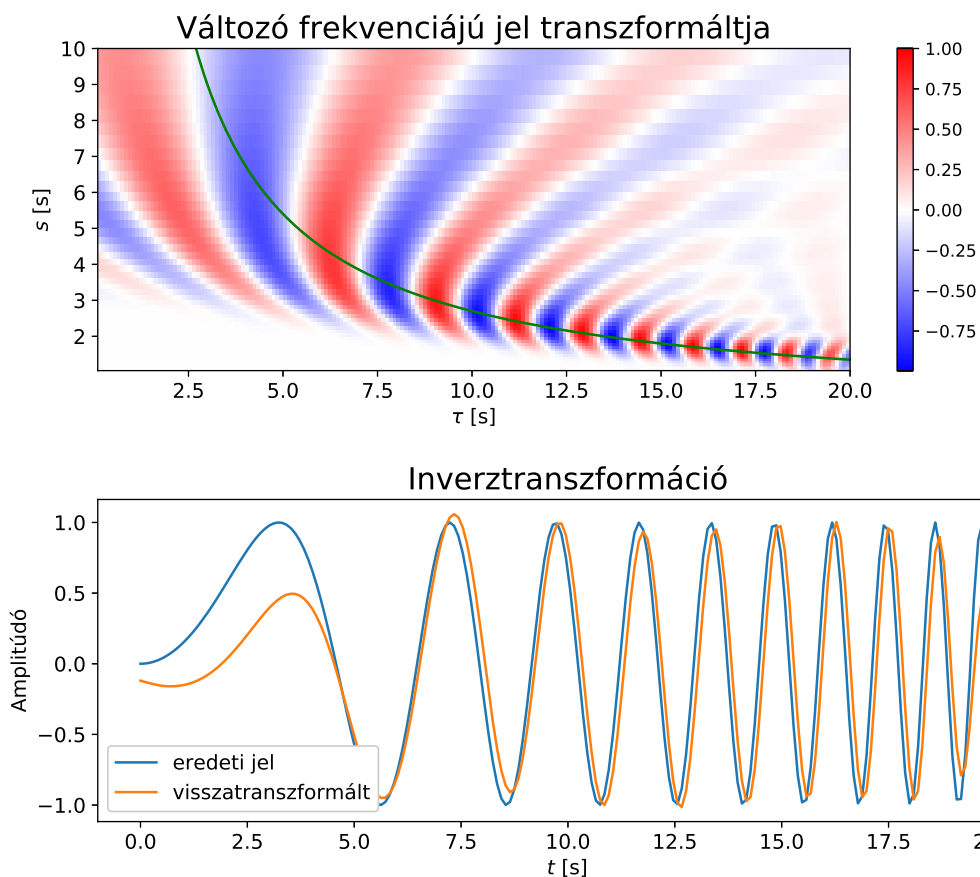
Az jel visszatranszformálásánál hasonló jelenséget tapasztalhatunk, mint az előző esetben: a visszatranszformált jel ugyan karakterisztikájában nagyban hasonlít az eredetihez, de némi eltérés van. Ez az eltérés a jel elején és végén a legnagyobb, ami arra utal, hogy a határoktól messze a wavelet transzformáció elég jól közelíti az elméleti eredményt. Mivel ezekben a példákban meglehetősen kevés pontban mintavételeztünk (mind s -ből, mind τ -ből, mind t -ből 10^{1-2} nagyságrendű adatunk van), ezért a "határokhoz közel eső tartomány" is relatíve nagy a teljes méretükhöz képest. Feltételezhetjük, hogy nagyobb adatoknál (10^{4-5} mintavételezés mellett) már a jel nagy részében az elméleti értékhez igen közel eső eredményt kapnánk.

3.3. Tanulságok az ismerkedés alapján

Azt tapasztaltuk, hogy eredményeink bár a várt eredményekhez hasonló jellegűek, elég nagy az eltérés azokhoz képest. Mivel a kód, amit használunk a lehetőségekhez képesti legegyszerűbb, ezért nem is várható el, hogy mindent pontosan visszaadjon, de tekintettel arra, hogy ebben a munkában nem a wavelet transzformáció minél pontosabb alkalma-



4. ábra. A (9) egyenlettel definiált, folytonosan változó frekvenciájú jel függvényalakja. Mivel a frekvencia időben nő, ezért a hullámok is egyre sűrűsödnek.



5. ábra. *fent*: Az időben változó frekvenciájú jel wavelet transzformáltja. Bár a domináns módusokra elég jól illik egy $s \sim \tau^{-1}$ görbe, a jelenlévő frekvenciák egyáltalán nem csak a várt görbe környékén vannak. *lent*: Az időben változó frekvenciájú jel, és transzformáltjának visszatranszformáltja. Bár a két jel itt is majdnem teljesen fedi egymást, a jel elején és végén jelentős az eltérés.

zása a cél, hanem a működésének és használatának megismerése, nem is fogunk mély elemzésekbe bonyolódni. A továbbiakban főként kvantitatív vizsgálatokat fogunk végezni, és ezeken keresztül igyekszünk minél jobban megérteni a wavelet transzformációt (és ennek a konkrét implementációnak) a működését.

4. D-dúr kánon szolmizálása

Johann Pachelbel, német barokk zeneszerző leghíresebb műve a D-dúr kánon. Fő jellegzetessége, hogy kánon alapdallama igen egyszerű, alapvető hangközök kombinációja és ismételtetése. Ebben a fejezetben a darab első motívumáról készült hangfelvételt fogjuk elemezni. [3] Ez egy kb. 4,4s-ig tartó részlet, amely az eredeti fájlban másodpercenkénti 44100 mintavételezéssel volt rögzítve. Annak érdekében, hogy a transzformációt gyorsan el tudjuk végezni, a mintavételezést lecsökkentettem 3675 Hz-re (ez ugyan azt is jelenti, hogy a magas frekvenciájú módusokat nem fogjuk tudni kimérni a mintavételi törvény miatt, de ahogy látni fogjuk így is elég információhoz juthatunk majd).

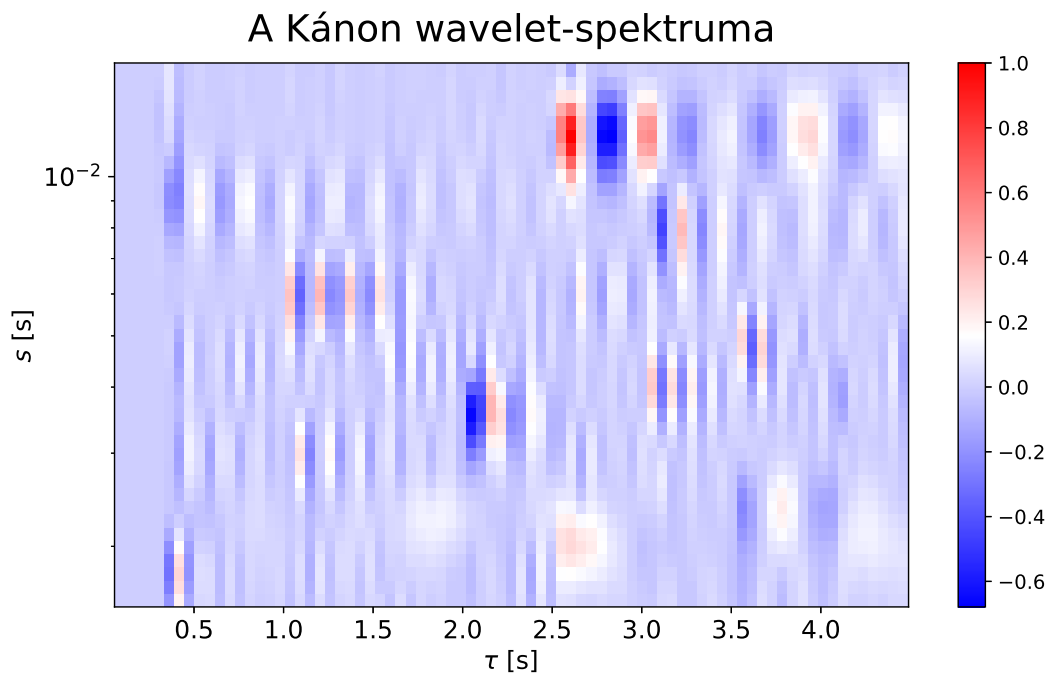
A minta transzformálása a csatolt `notebooks/canon.ipynb` fájlban található, a transzformált spektrumot pedig a 6. ábra ábrázolja. Látható, hogy a spektrumban kiemelkedő tartományok vannak, amelyek szabályos időközönként váltják egymást. Mivel a darabban egyenlő időközönként váltják egymást a hangok, pontosan ezt is vártuk.

Részletesebb elemzéshez, nézzük meg minden τ esetben, hogy melyik frekvenciamódus a legintenzívebb, ezeket a 7. ábrán fekete vonallal jelöltem meg. Ugyan a kapott mintázat meglehetősen zajos, de a háttér színezéséből, és az egymáshoz közel lévő vonalkákból kikövetkeztethető, hogy a spektrumnak melyek azok a tartományai, amelyek ténylegesen megszólaló hanghoz tartoznak. Mivel a második ütemben megszólaló első hang és a felhangjai olyan erősek, hogy elnyomja a többi, ezért az utolsó két hang nem azonosítható egyértelműen, de a dallam első hat hangja így is megtalálható. Ha két hang frekvenciáját összehasonlítjuk, akkor közöttük a zenei lépések száma a temperált hangskála alapján [4]:

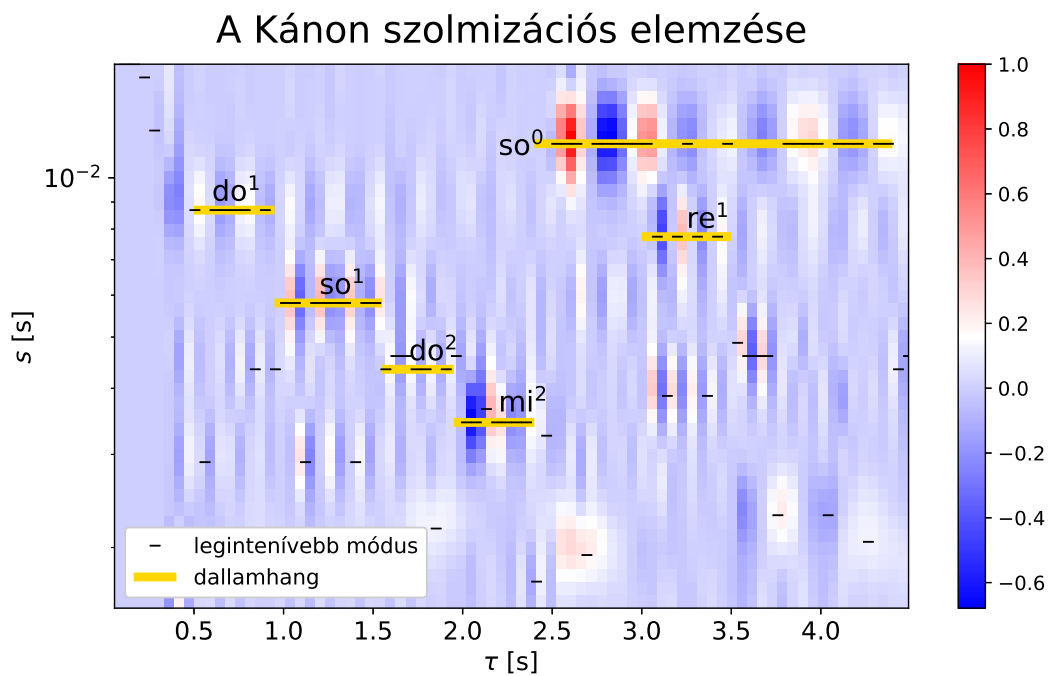
$$\Delta N = \log_{2^{1/12}} \left(\frac{s_1}{s_2} \right) \quad (10)$$

Ez alapján a megtalált hangok frekvenciáiból (illetve s -értékeiből) kiszámítottuk az első hangtól való távolságukat, amit az 1. táblázat tartalmaz.

Bár láttuk, hogy a transzformációt végző algoritmus elég kezdetleges, mégis meg tudtuk vele határozni egy zenedarab elejének szolmizációs hangjait. Ennek jelentősége nem csupán elméleti, hiszen pontosan ilyen elven működnek az elektronikus hangolók, a hang utáni kottázó programok és egyéb zenefeldolgozással kapcsolatos szoftverek.



6. ábra. A D-dúr kánon wavelet transzformáltja

7. ábra. A D-dúr kánon spektrumának elemzése. A fekete vonalak az adott τ -hoz tartozó legintenzívebb módust jelölik, a sárga sávok pedig ez alapján a dallamhangokat emelik ki.

n	s [10^{-2} s]	ΔN	hang
1	8,688	0	do ¹
2	5,799	7	so ¹
3	4,345	12	do ²
4	3,448	16	mi ²
5	11,60	-5	so ⁰
6	7,740	2	re ⁰

1. táblázat. A megtalált hangok s értékei, és ezek alapján zenei távolságuk az első hangtól. A zenei távolságot átkonvertálva szolmizációs hanggá, megkaphatjuk a zenedarab szolmizációját.

5. Dallam készítése spektrum alapján

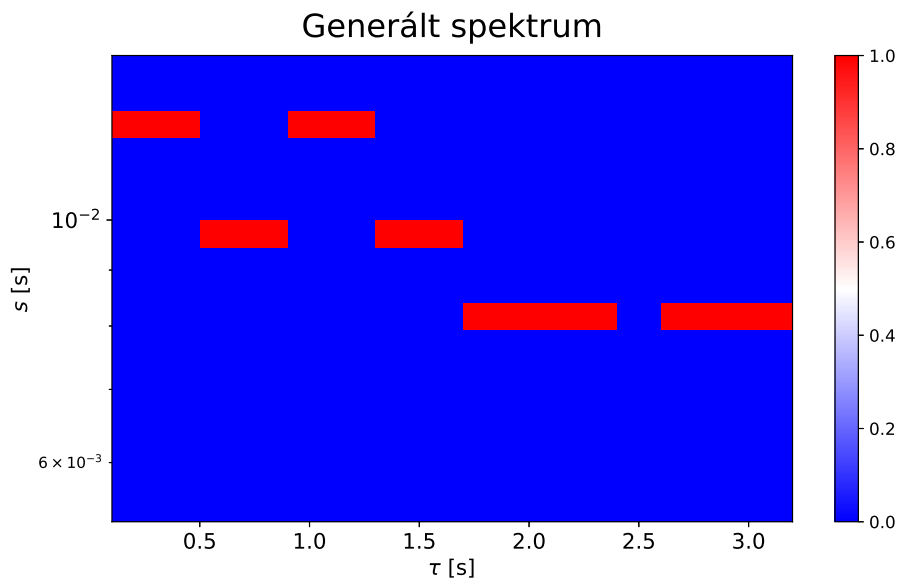
Az előző fejezetben egy dallam jelének transzformálásából nyertük ki a szolmizációját, most ennek a fordítottja a célunk: készítsük el a „dó-mi, dó-mi, szó, szó” dallam felvételét!

A feladat kivitelezése a csatolt `notebooks/inverse.ipynb` fájlban található. Első lépésben elkészítjük a spektrumot, ami a 8. ábrán látható. Ennek elkészítése úgy történt, hogy s felbontását úgy készítettem, hogy az egymást követő értékek egymás $2^{1/12}$ -szeresei legyenek, vagyis zeneileg egymást követő hangok. Ezután kiválasztottam egy alaphangot, és egyszerűen minden τ -ra értéket adtam ahhoz az s -hez tartozó cellához, amelyik az éppen megszólaló hangnak felel meg.

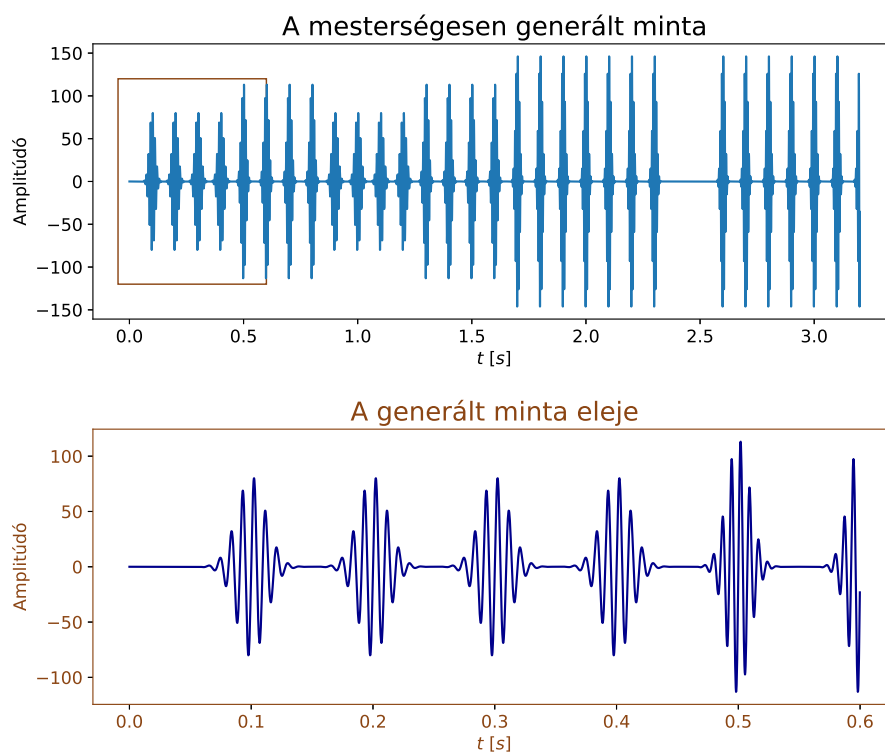
Erre a mintára alkalmaztam az inverz wavelet transzformációt. Az eredményt a 9. ábra ábrázolja.

Ha az inverz transzformáció eredményét kimentjük hangfájlba, és lejátszunk (a csatolt `files/melody.wav` fájl), akkor hallhatjuk, hogy elég furcsa dallamot kaptunk. Ugyan felismerhető a jellegzetes „dó-mi, dó-mi, szó, szó” dalocska, de ahelyett, hogy egyetlen hangot hallanánk, az adott hangot lüktetve halljuk (éles fülűek még azt is megszámlálhatják, hogy egy hang négy lüktetés idejéig tart). Ha megnézzük a 9. ábra alsó részén kinagyított részletet, láthatjuk, hogy a dallam valóban egymás utáni hullámcsomagokból áll.

Ennek magyarázata az, hogy kiinduló spektrum időfelbontása 32 darabból áll, vagyis egy hang 4 időegységig tart. Amikor erre alkalmazzuk az inverz transzformációt, akkor a konvolúció során az alapfüggvény négy jól elkülöníthető idődarabban találkozunk az adott móddal, aminek az lesz az következménye, hogy a kitranszformált jelben (amelynek időfelbontása sokkal nagyobb), maga az alapfüggvény fog megjelenni, négy egymástól jól elkülöníthető időpontban. Mivel azonban annyira nincsenek messze egymástól ezek a hullámcsomagok, ezért a fájl meghallgatása közben érzékeljük, hogy mekkora frekvenciával szól a hang, ezáltal a dallamot is felismerjük, de nem egy egyetlen hangot, hanem lüktetést hallunk.



8. ábra. A „dó-mi, dó-mi, szó, szó” dallam elkészített spektruma.



9. ábra. Az elkészített „dó-mi, dó-mi, szó, szó” dallam jele. *fent*: A kapott jel az inverz transzformáció után. *lent*: A minta első 0,6 s-a. Látható, hogy folytonos oszcillálás helyett négy hullámcsomag jelent meg.

A megoldást a problémára az jelentené, ha az eredeti spektrumot nagyobb időfelbontással készítettük volna el, ezáltal az egymás melletti hullámcsomagok nem lennének elkülönülve, hanem összefolynának, sőt kellően nagy felbontás esetén egy nagyjából egyenletes jellé állnának össze. Jelen esetben ezt azért nem alkalmaztuk, mert a használt algoritmus meglehetősen időigényes, így a lehetséges legkisebb, értelmezhető eredményt adó felbontást kerestük.

6. Összegzés

A dolgozatban a lehető legegyszerűbben implementált, de alapvető vizsgálatokra mégis alkalmas wavelet transzformációt végző algoritmus használatával ismerkedtünk. Láttuk, hogy a jelek oda- és visszatranszformálása meglehetősen torzítja az információt a diszkrét integrálás és az alacsony felbontás miatt, de még így is alkalmas a vizsgált minta egy-egy kvalitatív tulajdonságának vizsgálatára. Ezen kívül kipróbáltuk, hogyan lehet a wavelet transzformációval illetve az inverz wavelet transzformációval hangfájlokat készíteni, vagy zeneileg elemezni. Bár a kód egyszerűségéből fakadóan ezekben az esetekben is csak kezdetleges feladatokat végeztünk el, de bonyolultabb programok elvének az alapjaival ismerkedtünk meg.

Hivatkozások

- [1] Rubin H. Landau, Manuel J. Paez, Cristian C. Bordeianu. *A Survey of Computational Physics - Introductory Computational Science*. Princeton University Press, 2008.
- [2] A folytonos wavelet transzformáció angol nyelvű Wikipédia-szócikke [link](#)
- [3] Pachelbel D-dúr kánonjának felvétele [link](#)
- [4] Kanizsai Rita. *Zenei fogalmak és rendszerek a matematika nyelvén*. ELTE Numerikus Analízis Tanszék. 2012. [link](#)