A Fourier-analízis elmélete és gyakorlata

Nagy László
nagyl@inf.elte.hu

Témavezető:
Schipp Ferenc

2007. június 8.

Tartalomjegyzék

1. fejezet

Bevezető

A jelfeldolgozás során a bemenő jeleket frekvencia-komponensekre bontjuk. Ez a felbontás nem más, mint a bemenő jel (legyen az egy- vagy kétdimenziós, analóg, vagy digitális) valamilyen speciális függvénnyel történő szuperpozíciós felírása. Az ötlet természetesen nem új. Az első feljegyzések függvények közelítésére, függvény-együtthatók szuperpozíciójával a korai 1800-as évekre tehetőek, amikor *Joseph Fourier* felfedezte, hogy bizonyos tulajdonságokkal rendelkező függvények felbonthatóak szinusz- és koszinuszhullámok végtelen összegére. Ez matematikai áttörésnek számított gondolat számos, mára kialakult módszer alapját képezi. Joseph Fourier frekvencia-analízisbeli elmélete szerint (amit ma már csak Fourier-szintézis néven emlegetünk) bármely 2π szerint periodikus $f(x)$ függvény előállítható a következő hatvány-sor alakban:

$$f(x) = a_0 + \sum_{k=1}^{\infty}(a_k \cos kx + b_k \sin kx),$$

ahol az a_0, a_k és b_k együtthatókat az f függvényből határozhatjuk meg.

Fourier állítása kiemelkedően fontos szerepet játszott a matematikusok függvényekhez kötődő elméleteiben. A módszer egy teljesen más szemléletet nyújtott a függvények további vizsgálataihoz. Elméletével tulajdonképpen egy új, eddig még kinyitatlan kaput tárt szélesre a funkcionális világűr felé.

Néhány évtizedre rá, a tudósok – a szinusz és koszinusz hullámoknál – alkalmasabb függvényeket kezdtek el keresni, amelyek tartalmazzák a Fourier analízis alapjait és alkalmasak a különböző jelek közelítésére. Definíció szerint ugyanis e két függvény nem lokális, más szóval a tartójuk a valós számok $\mathbb{R}$ halmaza. 1807 után, megismerve a függvények új jelentését, a ourier-sorok konvergenciáját és az ortonormált rendszerek fogalmát, a matematikusok fokozatosan bevezették a *frekvencia analízis* mellett a *skála analízist*. Itt a függvényt különböző skálájú matematikai struktúrák segítségével kezdték el elemezni. E – mára nagyon fontos – módszer lényege, hogy vesznek egy függvényt, eltolják, átskálázzák, majd megvizsgálják, hogy a kapott struktúra előállítja-e az adott jelet. Ez utóbbi módszerrel a *wavelet analízis* foglalkozik.

A Fourier analízis azonban csak az 1960-as években válhatott a digitális jelfeldolgozás központi szerepévé. Ugyanis ez időtájt jelentek meg a transzformációt elvégző gyors, numerikusan jól kezelhető algoritmusok. Cooley és Tukey 1965-ben alkotta meg az első úgynevezett *FFT* algoritmust, amellyel a diszkrét Fourier-transzformáció kiszámítását gyakorlati feladatok esetére is megoldották, mintegy döntő lökést adva a digitális jelfeldolgozás további fejlődésének.

A dolgozat témája – mint ahogyan azt a cím is sugallja – a Fourier analízis eszköztárának részletes bemutatása. A leírás magában foglalja, a formálisabb matematikai leírás mellett, a módszer gyakorlati alkalmazását is. Tárgyaljuk a digitális jelfeldolgozás során alkalmazott fontosabb transzformációkat, melyek gyors elvégzésére bevezetjük a különböző FFT algoritmusokat. A dolgozat során bepillantunk a wavelet analízis varázslatos világába is. Bemutatva annak matematikai hátterét rávilágítunk a Fourier analízissel való fontos kapcsolatára. Végül a tárgyalt módszerek segítségével jelfeldolgozási alapokkal is megismerkedünk. A következőekben röviden áttekintjük az egyes fejezetek tartalmát.

Az 2. fejezetben áttekintjük a Fourier analízis alkalmazásához szükséges alapfogalmakat, definíciókat, fontos összefüggéseket. Mindezt tesszük a fejezet elején bevezetett általános csoporton, ez által lehetővé téve a későbbiek számára a speciális – számunkra fontos és gyakorlatban használt – csoportokra való leszűkítést. A jelfeldolgozásban fontos szerep játszó periodizáló operátor ismertetése után, a csoport karaktereinek segítségével, bevezetjük a Fourier-transzformáció (FT) és inverz Fourier-transzformáció (IFT) fogalmát. Mindeközben ismertetjük a transzformáció fontosabb operátorokkal való kapcsolatát, majd a bemutatott módszert kiterjesztjük a teljes L^2 térre.

A soron következő 3. fejezetben, a már megismert transzformációt, bemutatjuk néhány speciális csoportra. Ez által értelmezzük a Trigonometrikus Fourier-Transzformációt (TFT), a Trigonometrikus Fourier-Együtthatókat (TFE), a Trigonometrikus Fourier-Sort (TFS), a Diszkrét Fourier-Sort (DFS), valamint a Diszkrét Fourier-Transzformációt (DFT). A fejezet végén értelmezzük – a digitális jelfeldolgozás nélkülözhetetlen transzformációját – a z-transzformációt. A különböző konvergencia kritériumok mellett megmutatjuk a fontosabb operátorokkal való kapcsolatát. Végül leírjuk, hogy a számunkra fontos Fourier-transzformációkkal miként hozható kapcsolatba a z-transzformáció.

A 4. fejezetben a módszer gyors FFT algoritmusait vizsgáljuk meg. Először a DFT redundanciáit vizsgáljuk, majd ezek figyelembe vételével levezetjük az első gyorsított eljárást. Ezek után az algoritmusok általánosabb algebrai formáját adjuk meg, az előző szakaszban bevezetet műveletek általánosításával. Ezt követően a kialakított algoritmusainkat különböző szempotok szerint osztályozzuk. Végül bemutatjuk, hogy az FFT miként alkalmazható a jelfeldolgozás legfontosabb műveletének – a konvolúciónak – a gyors kiszámítására.

A rákövetkező 5. fejezet a wavletekről szól, pontosabban azok matematikai hátteréről. A bevezetés után rátérünk a wavelet analízis alapjául szolgáló multirezolúció ismertetésére, mely során definiáljuk a Riesz-bázis fogalmát. Ezt követően a Fourier-transzformáció segítségével bevezetjük a skálázási egyenletet, melynek megoldására egy példát is mutatunk. A fejezet végén ortonormált waveletek szerkesztésével is foglalkozunk, majd értelmezzük a Folytonos Wavelet-Transzformációt (CWT).

A 6. fejezetben a digitális jelfeldolgozás alapelemeivel ismerkedünk meg. Először az alul-, illetve felüláteresztő szűrők tulajdonságait mutatjuk be mind az idő-, mind a frekvencia-tartományon. Definiáljuk a skálázási egyenlethez nagyon hasonlító, wavelet egyenletet. Ismertetjük a szűrés két alapműveletét a downsampling és upsampling eljárásokat, továbbá bemutatjuk ezek hatását a különböző tartományokban. A bevezett szűrők és műveletek segítségével szűrőkészleteket építünk fel. Itt jegyezzük meg, hogy – a dolgozat terjedelmi korlátai miatt – csak 2-csatornás, FIR szűrőkészletek szerkesztésével foglalkotunk. A szűrőkészlet építésénél a tökéletes rekonstrukcióra törekszünk, ennek megkövetelése céljából mondunk ki feltételeket. Kapcsolatot teremtünk a diszkrét szűrők és a folytonos waveletek világa között, majd rekurziós módszereket mutatunk be a különböző szinteken lévő együtthatók előállítására, a keresett skála függvény közelítésére, valamint megmutatjuk, hogy az együtthatók ortogonalitásából miként tudunk következtetni a rendszerünk ortogonalitására. A diszkrét esethez tartozó Gyors Wavelet-Transzformáció (FWT) ismertetése mellett, bemutatunk néhány jelfeldolgozásbeli alkalmazást, mint például a zajmentesítés, vagy az éldetektálás folyamatát. Végül összehasonlítjuk a Fourier-transzformációt a waveletekkel, mely során rávilágítunk a Fourier analízis hézagira, hátrányaira.

Végül a 7. fejezetben a dolgozatban szereplő eredményeket értékeljük és foglaljuk össze, levonva a megfelelő következtetéseket.

Megjegyezzük, hogy a dolgozatban szereplő állítások, tételek bizonyítás nélkül szerepelnek, illetve csak azokat bizonyítjuk, melyeket elengedhetetlennek tartunk az elmélet teljes megértéséhez. Az egyes bizonyítások megtalálhatóak a jelzett forrásmunkákban.

2. fejezet

Matematikai háttér

2.1. Normált csoportok

A $(\mathbb{G}, +)$ kommutatív csoportból kiindulva bevezetjük a *normált csoport* fogalmát. Jelöljük 0-val a csoport nullelemét, $-x$-szel pedig az $x \in \mathbb{G}$ elem inverzét, ekkor, ha létezik a csoporton egy $x \to \|x\|$ leképezés (egy norma), akkor azt mondjuk, hogy a csoport normált. Ezen a csoporton bevezethető egy $\rho(x, y) := \|x - y\|$, $(x, y \in \mathbb{G})$ metrika, amelyben az $(x, y) \to x + y$ és $x \to -x$ leképezések folytonosak, ami annyit jelent, hogy az adott csoport *topologikus*. Könnyen ellenőrizhető, hogy a fenti metrika *transzláció invariáns*, azaz bármely $x, y \in \mathbb{G}$ elem távolsága megőrződik, ha mindkét elemet egy $z \in \mathbb{G}$ elemmel eltoljuk, nevezetesen a

$$\rho(x + z, y + z) = \rho(x, y) \quad (x, y, z) \in \mathbb{G} \tag{2.1}$$

egyenlőség teljesül, minden csoportbeli elemre. Akkor nevezzük a $(\mathbb{G}, \rho)$ metrikus teret *lokálisan kompaktnak*, ha a $(\mathbb{G}, +)$ normált csoport zárt gömbjei kompakt halmazok. Amennyiben a $(\mathbb{G}, \rho)$ metrikus tér kompakt, abban az esetben a $\mathbb{G}$ *normált csoport kompakt*, továbbá, ha az $inf\{\|x\| : x \neq 0\} > 0$ egyenlőtlenség is fenn áll, úgy a szóban forgó *csoport diszkrét*.

A következőkben – ezen új fogalmak segítségével – bevezetünk néhány speciális csoportot. Ismeretes ugyanis, hogy a valós számok $\mathbb{R}$ halmaza a $+$ – szokásos – összeadásra nézve Abel-csoport és az $\|x\| := |x|$, $(x \in \mathbb{R})$ norma ezen a csoporton, ami azt jelenti, hogy az $(\mathbb{R}, +)$ normált csoportot alkot. Az egész számok $\mathbb{Z}$ halmaza pedig – ugyanezzel a normával – az $\mathbb{R}$-nek egy normált, diszkrét részcsoportja. A csoportelméletben bevezetett faktorcsoport [11] definíciójából következik, hogy az $\mathbb{R}/\mathbb{Z}$ (faktorcsoport) izomorf a $[0, 1)$ intervallum modulo 1 összeadással vett csoportjával, amit a továbbiakban

$\mathbb{M}$-mel jelölünk. Bevezetve a $\dot{+}$ jelölést a csoportműveletre, a

$$x\dot{+}y := \begin{cases} x+y, & \text{ha } x+y<1, \\ x+y-1, & \text{ha } x+y\geq 1 \end{cases} \tag{2.2}$$

formula áll fenn minden $x, y \in \mathbb{M} := [0,1)$ esetén. Továbbá belátható, hogy a csoport nulleleme a 0 szám, az $x \in \mathbb{M}$ elem inverze pedig az $1-x$ és az $\|x\| := min\{x, 1-x\}$ $(x \in \mathbb{M})$ normával az $(\mathbb{M}, \dot{+})$ csoport is normált. Ennek segítségével az

$$\mathbb{M}_m := \left\{ \frac{k}{m} : k = 0, 1, \ldots, m-1 \right\} \quad (\forall m \in \mathbb{N}^*) \tag{2.3}$$

definícióval megadott csoport az $\mathbb{M}$ egy diszkrét m-edrendű ciklikus részcsoportja. Az *egydimenziós tórusz*, másnéven az egység abszolút értékű komplex számok

$$\mathbb{T} := \{z \in \mathbb{C} : \|z\| = 1\} \tag{2.4}$$

halmaza a szorzásra nézve csoportot alkot, amelynek nulleleme az 1 szám, a $z \in \mathbb{T}$ elem inverze pedig a z szám $\bar{z}$ komplex konjugáltja. A *komplex trigonometrikus függvény*, tehát az

$$\epsilon(t) := \exp(i2\pi t) = \cos(2\pi t) + i\sin(2\pi t) \quad (t \in \mathbb{R}) \tag{2.5}$$

formulával definiált leképezés, $\mathbb{M}$-re vonatkozó leszűkítése egy izomorfizmus $\mathbb{M}$ és $\mathbb{T}$ között, amely az $\mathbb{M}$-beli elemek $\dot{+}$ összeadását az elemek $\mathbb{T}$-beli képeinek a szorzásába viszi át, azaz ϵ az $(\mathbb{M}, \dot{+})$ *csoport karaktere* (a karaktereket részletesebben az (2.3) szakaszban mutatjuk be). Végül a $\|z\| := |1-z|$ $(z \in \mathbb{T})$ leképezéssel a $(\mathbb{T}, \cdot)$ csoport is normált.

Az eltolás műveletével természetes módon értelmezhetjük függvények körében a *transzláció operátorát.* Tetszőleges $\mathbb{G}$ csoporton értelmezett f komplex értékű függvényre legyen

$$(\tau_s f)(x) := f(x+s) \quad (s, x \in \mathbb{G}). \tag{2.6}$$

Látható, hogy a trsanszláció nem más, mint a függvény argumentumának – adott csoportbeli elemmel való – eltolása.

2.2. Periodizáló operátor

A jelfeldolgozásban – és persze a további vizsgálatainkban – fontos szerepet játszanak azok az operátorok, melyek egy adott függvényhez annak periodikus változatát rendelik. Jelölje L^p_T $(T > 0,\ 1 \leq p \leq \infty)$ azoknak az $\mathbb{R}$-en értelmezett, lokálisan integrálható, T-szerint periodikus függvények halmazát, melyeknek a $[0,T)$ intervallumra való leszükítése $L^p[0,T]$ térbeli. Nyilvánvaló, hogy ezekre a terekre $L_T^p \subset L_{2T}^p$ tartalmazás teljesül. Legyen $f \in L^1(\mathbb{R})$ ekkor az

$$(E_T f)(x) := \sum_{k\in\mathbb{Z}} f(x+kT) = \sum_{k\in\mathbb{Z}} (\tau_{kT} f)(x)(x \in \mathbb{R}) \tag{2.7}$$

utasítással értelmezett operátort *periodizáló operátornak* nevezzük. Mivel

$$\sum_{k\in\mathbb{Z}} \int_0^T |f(x+kT)|\ dx = \int_{-\infty}^{\infty} |f(x)|\ dx < \infty,$$

ezért a (2.7) sor majdnem minden $x \in \mathbb{R}$ pontban abszolút konvergens, azaz a definíció korrekt. Továbbá a fent bevezetett $E_T : L^1(\mathbb{R}) \to L_T^1$ leképezés egy korlátos líneárs operátor, amelyre

$$\int_0^T (E_T f)(x)\ dx = \int_{-\infty}^{\infty} f(x)\ dx, \quad \int_0^T |(E_T f)(x)|\ dx \le \int_{-\infty}^{\infty} |f(x)|\ dx.$$

Bizonyítható, hogy ez az operátor nemcsak számokra, hanem T-szerint periodikus függvényekre nézve is homogén, azaz bármely λ, T-szerint periodikus, Lebesgue-mérhető függvény esetén, ha $f, \lambda f \in L^1(\mathbb{R})$, akkor

$$E_T(\lambda f) = \lambda E_T(f).$$

Azonban a most bevezetett periodizáló operátornak nemcsak a fent leírt interpretációja van. Ugyanis E_T felfogható egy speciális *várhatóérték* operátornak, amely nagyon sok vonatkozásban hasonlít az integrál funkcionálhoz, melynek következményeképpen az ott bevezetett – integrállal kapcsolatos – alapvető fogalmak átvihetők erre az operátorra. Ez által lehetőség nyílik a skaláris szorzat és az ortogonalitás fogalmának általánosítására, kiindula az

$$(f,g) \to E_T(f\bar{g}) \quad (f,g \in L^2(\mathbb{R}))$$

bilineáris operátorokból. Akkor mondjuk, hogy az $f,g \in L^2(\mathbb{R})$ függvények *E_T-ortogonálisak*, ha

$$E_T(f\bar{g}) = 0.$$

Ez a definíció valóban tekinthető általánosításnak, ugyanis az

$$\int_0^T E_T(f\bar{g})(x)\ dx = \int_{\mathbb{R}} f(x)\overline{g(x)}\ dx =< f,g >, \quad (f,g \in L^2(\mathbb{R}))$$

egyenlőségek miatt az E_T-ortogonalitásból következik a szokásos ortogonalitás.

2.3. Karakterek

Következő lépésként értelmezzük a korábban már bevezetett $(\mathbb{G}, +)$ normált csoport karaktereit, a komplex trigonometrikus függvények tulajdonságai alapján.

Akkor nevezünk egy $\gamma : \mathbb{G} \to \mathbb{T}$ leképezést a *csoport karakterének (azaz trigonometrikus függvényének)*, ha γ folytonos és eleget tesz a

$$\gamma(x+y) = \gamma(x)\gamma(y) \quad (x, y \in \mathbb{G}) \tag{2.8}$$

függvényegyenletnek. Ez a feltétel azt jelenti, hogy a szóban forgó leképezés egy homomorfizmus $\mathbb{G}$ és a komplex tórusz között. $\mathbb{G}$ karaktereinek a halmazát $\widehat{\mathbb{G}}$ szimbólummal fogjuk jelölni. Nyilvánvalóan – minthogy két karakter szorzata is karakter – $\widehat{\mathbb{G}}$ a $\gamma_0(x) := 1$ $(x \in \mathbb{G})$ egységelemmel csoportot alkot, ahol a $\gamma \in \widehat{\mathbb{G}}$ karakter inverzét annak komplex $\bar{\gamma} = \frac{1}{\gamma}$ konjugáltja szolgáltatja. Az így kapott $(\widehat{\mathbb{G}}, \cdot)$ csoportot a $(\mathbb{G}, +)$ *csoport duálisának* nevezzük.

Csoportok direktszorzatát felhaszálva a karakterek kiterjeszthetők n dimenzióra, melyekre igaz, hogy a duális csoportbeli karakterek előállíthatók a csoportok karaktereinek Kronecker-szorzataként:

$$\gamma(x) = (\gamma_1 \times \gamma_2 \times \cdots \times \gamma_n)(x) := \gamma_1(x_1)\gamma_2(x_2)\cdots\gamma_n(x_n), \tag{2.9}$$

$$(x = (x_1, \ldots, x_n) \in \mathbb{G})$$

Ennek megfelelően, $\mathbb{G}$-vel jelölve az $(\mathbb{R}^n, +)$, $(\mathbb{Z}^n, +)$ és $(\mathbb{M}^n, \dot{+})$ csoportok bármelyikét, a csoport duálisának karakterei egységes formában írhatók fel

$$\epsilon_s(t) := exp(2\pi i\langle s, t\rangle) \quad (t \in \mathbb{G}, s \in \widehat{\mathbb{G}})$$

alakban, ahol $\langle s, t\rangle$ a szokásos $\mathbb{R}^n$-beli skaláris szorzatot jelöli.

Hasonlóan a komplex trigonometrikus rendszerhez, a kompakt csoportok karakterei is ortonormált rendszert alkotnak a normált Haar-mértékkel [3] $L^2_m(\mathbb{G})$ térben, azaz μ-vel jelölve a $\mathbb{G}$ csoport normált Haar-mértékét teljesül a

$$\int_{\mathbb{G}} \gamma_1(x)\overline{\gamma_2(x)}\, d\mu(x) = \delta_{\gamma_1,\gamma_2} \quad (\gamma_1, \gamma_2 \in \widehat{\mathbb{G}}) \tag{2.10}$$

egyenlet, ahol δ_{rs} a Kronecker-féle szimbólum.

2.4. Fourier-transzformált

Ebben a pontban – az imént ismertetett karakterek segítségével – definiáljuk a $\mathbb{G}$-n értelmezett, a $\mathbb{G}$ csoport μ Haar-mértéke szerint integrálható függvények Fourier-transzformáltját. Ezek után bemutatjuk a – nemsokára bevezetésre kerülő – moduláció, dilatáció és – a már definiált – transzláció operátorokkal való kapcsolatát. Legvégül definiálunk egy olyan, függvények közötti műveletet, amely nem vezet ki az $L^1_m(\mathbb{G})$ térből.

Tehát legyen $f \in L^1_\mu(\mathbb{G})$, ekkor a $\widehat{\mathbb{G}}$ halmazon értelmezett

$$(\mathcal{F}f)(\gamma) := \widehat{f}(\gamma) := \int_{\mathbb{G}} f(t)\overline{\gamma(t)}\, d\mu(t) \quad (\gamma \in \widehat{\mathbb{G}}) \tag{2.11}$$

$\hat{f}$ függvényt az f függvény *Fourier-transzformáltjának* nevezzük. Speciálisan, a már bevezetett csoportok

$$\epsilon_x(s) = \epsilon_{x_1}(s_1) \cdots \epsilon_{x_n}(s_n) \quad (x = (x_1, \ldots, x_n) \in \widehat{\mathbb{G}}^n, s = (s_1, \ldots, s_n) \in \mathbb{G}^n)$$

karaktereinek alkalmazásával az $f \in L^1_\mu(\mathbb{G})$ függvény Fourier-transzfolmáltja a

$$(\mathcal{F}f)(x) := \widehat{f}(x) := \int_{\mathbb{G}} f(t)\overline{\epsilon_x(t)}\, d\mu(t) \quad (x \in \widehat{\mathbb{G}}) \tag{2.12}$$

alakban írható fel, ahol μ a $\mathbb{G} \in \{\mathbb{R}, \mathbb{M}\}$ esetben a Lebesgue-mértéket, míg $\mathbb{G} \in \{\mathbb{Z}, \mathbb{M}_m\}$ esetben pedig a diszkrét mértéket jelenti. A 3. fejezetben részletesen bemutatjuk az egyes speciális csoportokon értelmezett transzformációkat.

A most bevezetett Fourier-transzformált fontos tulajdonsága, hogy bármely $f \in L^1_m(\mathbb{G})$ függvény $\mathcal{F}f$ Fourier-transzformáltja a $\widehat{\mathbb{G}}$ halmazon értelmezett, a végtelenben eltűnő függvény, továbbá $\mathcal{F} : L^1_m(\mathbb{G}) \to \mathcal{C}_0(\widehat{\mathbb{G}})$ korlátos, líneáris operátor, amelyre $\|\mathcal{F}f\|_\infty \leq \|f\|_1$ $(f \in L^1_m(\mathbb{G}))$ teljesül [3].

A harmónikus analízisben fontos szerep jut bizonyos operátoroknak, nevezetesen az argumentum eltolásaival és a karakterek szorzásával összefüggő transzformációkból származtatott *dilatációnak*

$$(\delta_z f)(s) := f(zs) \quad (z \in \mathbb{R}^*, s \in \mathbb{R}) \tag{2.13}$$

és *modulációnak*

$$(\nu_y f)(s) := \epsilon_y(s) f(s) \quad (s \in \mathbb{G}, y \in \widehat{\mathbb{G}}). \tag{2.14}$$

Ezen, valamint az 1. fejezetben bevezetett, *transzláció* operátorok szoros kapcsolatban állnak a Fourier-transzformációval, nevezetesen

$$\mathcal{F}(\tau_a f) = \nu_a \mathcal{F}f \quad (a \in \mathbb{G}), \tag{2.15}$$

$$\mathcal{F}(\nu_b f) = \tau_{-b} \mathcal{F}f \quad (b \in \widehat{\mathbb{G}}), \tag{2.16}$$

$$\mathcal{F}(\delta_c f) = c^{-1}\delta_{c^{-1}}\mathcal{F}f \quad (\mathbb{G} = \mathbb{R}, c > 0). \tag{2.17}$$

Tehát, mint látható, a bevezetett operátorok a Fourier-transzformációval nem cserélhetők fel, pontosabban a felcserélés csak a fenti szabályok alkalmazásával lehetséges.

Köztudott, hogy az $L^1_m(\mathbb{R})$ líneáris tér nem zárt a függvények pontonkénti szorzására nézve. Az említett – sőt, általánosan az $L^1_m(\mathbb{G})$ – téren bevezethető, egy ettől eltérő bínáris művelet, az úgynevezett *konvolúció*, amely már nem vezet ki a térből. Ez a transzformáció szoros kapcsolatban van a $(\mathbb{G}, +)$ csoport struktúrájával és a rajta értelmezett Fourier-transzformáltjával. Minthogy bármely $f, g \in L^1_m(\mathbb{G})$ függvényre az

$$(f * g)(x) := \int_{\mathbb{G}} f(t)g(x-t)\, d\mu(t) \quad (x \in \mathbb{G}) \tag{2.18}$$

integrál minden $(x \in \mathbb{G})$ pontban létezik és véges μ majdnem mindenütt, ezért a most – (2.18) alatt – értelmezett $(f * g) \in L^1_m(\mathbb{G})$ függvényt az $f, g \in L^1_m(\mathbb{G})$ függvények *konvolúciójának* nevezzük. Természetesen a konvolúció és Fourier-transzformáció között is szoros, sőt, igen fontos kapcsolat van, nevezetesen bármely két $f, g \in L^1_m(\mathbb{G})$ függvényre az alábbi

$$\widehat{f * g} = \widehat{f}\widehat{g} \tag{2.19}$$

egyenlőség teljesül [3]. Ezen – (2.19) alatti – azonosság interpretációja az, hogy a Fourier-transzformált a függvények konvolócióját a transzformált függvények szorzatába viszi át. A konvolúció kiterjeszthető az $X \in \{L^p_m(\mathbb{G}), C(\mathbb{G}), C_0(\mathbb{G})\}$ $(1 \leq p \leq \infty)$ terekre, melyekre fennáll, hogy bármely $g \in L^1_m(\mathbb{G})$ és $f \in X$ függvények esetén létezik az $f * g$ konvolúció, melyre $f * g \in X$ és

$$\|f * g\| \leq \|g\|_1 \|f\|_X \quad (g \in L^1_m(\mathbb{G}), f \in X). \tag{2.20}$$

2.5. Inverziós formula

Most vizsgáljuk meg, hogyan, milyen feltételek mellett lehet az integrálható függvényeket visszaállítani azok Fourier-transzformáltjából. Ezúttal olyan integrálható függvényekből indulunk ki, amelynek Fourier-transzformáltja is integrálható az adott csoport duálisán, amely integrálhatóság az ℓ^1 tér esetén a Fourier-együtthatókból alkotott sor abszolút konvergenciájával ekvivalens. A formulával összefüggésben bevezetjük a *Fourier-transzformált adjungáltját*.

Kiindulva az

$$\langle f, g \rangle := \int_{\mathbb{G}} f(t)\bar{g}(t)\, d\mu(t) \tag{2.21}$$

$$\left(f \in L^p_m(\mathbb{G}), g \in L^q_m(\mathbb{G}), \frac{1}{p} + \frac{1}{q} = 1, 1 \le p, q \le \infty\right)$$

funkcionálból, az $\mathcal{F}$-nek erre vonatkozó *adjungáltja* az

$$(\mathcal{F}^* f)(x) := (\mathcal{F} f)(-x) \quad (x \in \widehat{\mathbb{G}}) \tag{2.22}$$

operátor, amire

$$\langle \mathcal{F} f, g \rangle = \langle f, \mathcal{F}^* g \rangle \quad (f, g \in L^1(\mathbb{R})) \tag{2.23}$$

teljesül. Ezen operátor segítségével felírható az *inverziós formula* általános alakja [3], azaz, ha $f \in L^1_m(\mathbb{G})$ és $g \in L^1_m(\widehat{\mathbb{G}})$, akkor

$$\mathcal{F}^*(\mathcal{F} f) = f. \tag{2.24}$$

A 3. fejezetben bevezetésre kerülő speciális Fourier-transzformációk bemutatásakor kitérünk az – adott transzformációhoz tartozó – inverziós formulák pontos alakjára.

2.6. Transzformáció kiterjesztése

Ebben a részben $\mathbb{G} = \mathbb{R}$ esetre szorítkozunk, mely esetében a Fourier-transzformáció (2.12) alapján integrálható függvényekre értelmezhető. A definíciót kiterjesztjük $L^2(\mathbb{R})$-beli függvényekre, amely téren a kiterjesztett operátor *unitér*, azaz $\mathcal{F}$-fel jelölve a kiterjesztett operátort

$$\langle \mathcal{F} f, \mathcal{F} g \rangle = \langle f, g \rangle \quad (f, g \in L^2(\mathbb{R})) \tag{2.25}$$

teljesül.

Mivel az (2.25) egyenlőség fennáll az úgynevezett *diadikus intervallumok karakterisztikus függvényeire*, azaz az

$$\mathcal{I} := \{\chi_I : I = [k2^n, (k+1)2^n), (k, n \in \mathbb{Z})\} \subset L^1(\mathbb{R}) \tag{2.26}$$

függvényosztály elemeire (nevezetesen bármely $I, J \in \mathcal{I}$ esetén $\langle \mathcal{F}\chi_I, \mathcal{F}\chi_J \rangle = \langle \chi_I, \chi_J \rangle$), továbbá az $\mathcal{I}$ zárt rendszer az egész $L^2(\mathbb{R})$ térben, ezért a Fourier-transzformáció az (2.25) tulajdonság megtartásával kiterjeszthető az egész $L^2(\mathbb{R})$ térre. Ennek alapján az $L^1(\mathbb{R}) \bigcup L^2(\mathbb{R})$ halmazra kiterjesztett Fourier-transzformáltat is $\mathcal{F}$-fel fogjuk jelölni.

A korábban bevezetett $\mathcal{F}^*$ operátor a következő, $(\mathcal{F}^* f)(x) := (\mathcal{F} f)(-x)$ utasítással ($x \in \mathbb{R}, f \in L^2(\mathbb{R})$) kiterjeszthető az $L^2(\mathbb{R})$ térre és a definíció, valamint (2.25) alapján

$$\|\mathcal{F}^* f\|_2 = \|\mathcal{F} f\|_2 = \|f\|_2 \quad (f \in L^2(\mathbb{R})).$$

Továbbá bármely $f, g \in L^2(\mathbb{R})$ függvényre az (2.23) alatti azonosság most is teljesül, azaz $\mathcal{F}^*$ a *Fourier-transzformált adjungáltja.*

Végül bármely $f, g \in L^2(\mathbb{R})$ függvényre $\langle \mathcal{F}(\mathcal{F}^* f), g\rangle = \langle f, g\rangle$ miatt $\mathcal{F}(\mathcal{F}^* f) = f$ $(f \in L^2(\mathbb{R}))$, tehát $\mathcal{F}$ értékkészkete az $L^2(\mathbb{R})$ tér [3]. Következésképpen

$$\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$$

olyan *bijekció, melynek inverze* $\mathcal{F}^*$, azaz

$$\mathcal{F}^{-1} = \mathcal{F}^*.$$

A konvolúció és a Fourier-transzformáció közötti kapcsolat a kiterjesztett $\mathcal{F}$ operátorra abban formában áll fent, hogy ha $f \in L^1(\mathbb{R})$ és $g \in L^2(\mathbb{R})$, akkor a

$$\mathcal{F}(f * g) = \mathcal{F}f\mathcal{F}g$$

állítás teljesül.

3. fejezet

Transzformációk

3.1. Trigonometrikus Fourier-Transzformáció

Mint ahogyan az a bevezetőben olvasható volt, a Fourier-transzformáció egy leképezés az időtartományról ($\mathbb{G}$) a frekvenciatartományra ($\widehat{\mathbb{G}}$). Az előző fejezetben bevezetett műveletek a legszebben akkor néznek ki, ha mind az idő-, mind a frekvencia-tartomány a valós számok halmaza, azaz $\mathbb{G} = \widehat{\mathbb{G}} = \mathbb{R}$. Ebben a szakaszban ezt az esetet tárgyaljuk, bemutatva az általános definíciók ide vonatkozó alakját.

Elsőként vizsgáljuk meg a (2.8) alatt bezetett karaktereket. A korábban már említett

$$\epsilon_t(x) = exp(2\pi itx) \ (t, x \in \mathbb{R}) \tag{3.1}$$

függvények bármely $t \in \mathbb{R}$ esetén az $(\mathbb{R}, +)$ csoport karakterei, továbbá bebizonyítható, hogy a $t \to \epsilon_t$ leképezés egy izomorfizmus, melynek következményeképpen a szóban forgó csoport duálisa önmaga. Az állítás a (2.9) alatt bevezett n-dimenziós esetre is érvényes, tehát az $(\mathbb{R}^n, +)$ karakter csoportja izomorf önmagával.

A (2.11) alatt bevezetett Fourier-transzformáció $\mathbb{G} = \mathbb{R}$ esetre vonatkozó megfelelője az $L^1(\mathbb{R})$-beli függvények

$$\widehat{f}(x) = \int_{\mathbb{R}} f(t)exp(-2\pi ixt)\ dt \quad (f \in L^1(\mathbb{R}), x \in \mathbb{R}) \tag{3.2}$$

úgynevezett *Trigonometrikus Fourier-Transzformáltja* (TFT).

A (2.18) alatt bevezetett konvolúció TFT-re vontakozó alakja az alábbi formában írható fel:

$$(f * g)(x) := \int_{-\infty}^{\infty} f(t)g(x-t)\ dt \quad (x \in \mathbb{G} = \mathbb{R}, f, g \in L^1(\mathbb{R})) \tag{3.3}$$

és ezt viszi át a Fourier-transzformáció a transzformáltak szorzatába a (2.19) egyenlőségnek megfelelően.

Az (2.24) inverziós formula $\mathbb{G} = \mathbb{R}$ csoportra vonatkozó alakja az

$$f(x) := \int_{-\infty}^{\infty} \widehat{f}(t)\epsilon_x(t)\ dt\ (f, \widehat{f} \in \mathbb{R}) \tag{3.4}$$

alakot ölti. Ezt, a (3.2) alatt definiált transzfomációval összevetve látható, hogy a két művelet egy konstans szorzó erejéig megegyezik, amiből látszik, hogy a két tartomány ($\mathbb{R}$) valóban izomorfak egymással.

Az előző fejezet végén azt is megmutattuk, hogy a transzformáció kiterjeszthető az $L^2(\mathbb{R})$ térre unitér operátorrá. A (2.25) egyenlőséget $g = f$ esetre alkalmazva az *energiatételhez*, vagy más néven *Parseval-egyenlőséghez* jutunk:

$$\int_{-\infty}^{\infty} |f(t)|^2\ dt = \int_{-\infty}^{\infty} |\mathcal{F}f(t)|^2\ dt. \tag{3.5}$$

Még egy fontos összefüggést is megemlítünk, amely a mintavételezésre vonatkozik. Nevezetesen, ha egy diszkrét g függvényt (azaz sorozatot) úgy állítunk elő egy folytonos idejű f függvényből, hogy azt az egész időpontokban mintavételezzük, akkor a g Fourier-transzformáltja előállítható f Fourier-transzformáltjának periodizáltjaként. Alkalmazva a (2.2) alatt bevezetett periodizáló operátor jelöléseit és bevezetve az $E_T := E$ (ha $T = 1$) jelölést formálisan a

$$\widehat{g}(x) = (E\widehat{f})(x)$$

alakhoz jutunk. Ez által elméleti és gyakorlati szempontból egyaránt fontos kapcsolatot teremtettünk a folytonos idejű Fourier-integrál és a diszkrét sorozatokkal foglalkozó DFT reprezentáció között.

3.2. Trigonometrikus Fourier-Együtthatók

Az előző esethez hasonlóan most is újratárgyaljuk a bevezetett definíciókat, azonban most a transzformáció értelmezési-, azaz idő-tartománya $\mathbb{G} = \mathbb{M}$.

Kezdve most is a karakterekkel, megmutatható, hogy az ϵ_n $(n \in \mathbb{Z})$ 1-szerint periódikus trigonometrikus függvények karakterei az $(\mathbb{M}, \dot{+})$ csoportnak és az $n \to \epsilon_n$ leképezés egy izomorfizmus a $\mathbb{Z}$ és az említett csoportok között, tehát az $(\mathbb{M}, \dot{+})$ csoport duálisa $(\mathbb{Z}, +)$, azaz $\widehat{\mathbb{G}} = \mathbb{Z}$. (2.9) alkalmazásával, n-dimenziós esetben a $(\mathbb{M}^n, \dot{+})$ karakter csoportja $(\mathbb{Z}^n, +)$ csoporttal azonosítható.

A (2.11) alatt bevezetett Fourier-transzformáció $\mathbb{G} = \mathbb{M}$ esetre vonatkozó megfelelője az $L^1(\mathbb{M})$-beli függvények

$$\widehat{f}(n) = \int_{\mathbb{M}} f(t) \exp(-2\pi int)\ dt \quad (f \in L^1(\mathbb{M}), n \in \mathbb{Z}) \tag{3.6}$$

úgynevezett *Trigonometrikus Fourier-Együtthatója* (TFE). Látható tehát, hogy a transzformáció adott integrálható függvényből, a függvényhez tartozó együtthatók sorozatát állítja elő.

A (2.18) alatt bevezetett konvolúció TFE-re vontakozó alakja az alábbi formában írható fel:

$$(f * g)(x) := \int_0^1 f(t) g(x-t)\ dt \quad (x \in \mathbb{G} = \mathbb{M}, f, g \in L^1(\mathbb{M})). \tag{3.7}$$

Természetesen a (3.6) transzformáció a most bevezetett konvolúciót is a transzformáltak szorzatába viszi át.

Az (2.24) inverziós formula $\mathbb{G} = \mathbb{M}$ csoportra vonatkozó alakja az

$$f(x) := \sum_{n \in \mathbb{Z}} \widehat{f}(n) \epsilon_n(x) \quad (f \in L^1(\mathbb{M}), \widehat{f} \in \ell^1(\mathbb{Z})) \tag{3.8}$$

alakot ölti, amely alatt a (3.6) transzformációval előállított $\hat{f}$ együtthatókból a függvény előállítását értjük.

Egy $\mathbb{M} = [0, 1)$ intervallumon értelmezett függvény, tulajdonképpen felfoghatók egy periodikus – nem feltétlenül folytonos – függvény egyetlen periódusának, ahol a periódusidőt az intervallum hossza – jelen esetben 1 – jelöli. Ez által az $f \in \mathbb{M}$ függvények kiterjeszthetők az egész valós számegyenesre periodikus függvényekké, jelöljük ezeket $\widetilde{f}$ szimbólummal. Általánosabb formában való felíráshoz jelöljük T_0-lal a periodusidőt.

Egy periodikus, folytonos idejű $\widetilde{f}(t)$ függvény ismeretesen végtelen sok diszkrét komponensből álló Fourier-sorba fejthető, az

$$\widetilde{f}(t) = \sum_{m=-\infty}^{\infty} c_m e^{im\omega_0 t} \tag{3.9}$$

kifejtéssel, ahol a c_m komplex Fourier-együtthatók tetszőleges m egész számra a

$$c_m = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} \widetilde{f}(t) e^{-im\omega_0 t}\ dt \tag{3.10}$$

egy periódusra vett integrálból számolhatók. Az $\widetilde{f}(t)$ függvény T_0 periódusideje és a Fourier-sorfejtés $m\omega_0$ frekvenciájú tagjai közötti kapcsolat az ω_0 alapharmonikuson keresztül az $\omega_0 = 2\pi/T_0$ egyenletből következik.

3.3. Trigonometrikus Fourier-Sor

Ebben az esetben az idő-tartomány az egész számok halmaza, azaz $\mathbb{G} = \mathbb{Z}$. Ennek következménye, hogy a bevezetésre kerülő transzformáció $\ell^1(\mathbb{Z})$-beli konvergens sorozatok felett operál.

Előzőekben – (3.2) alatt – megmutattuk, hogy az $(\mathbb{M}, \dot{+})$ csoport duálisa $(\mathbb{Z}, +)$. Minthogy az ϵ_t $(t \in \mathbb{M})$ függvények a $(\mathbb{Z}, +)$ csoport karakterei, és valamennyi karakter megkapható ilyen módon, továbbá az

$$\epsilon_{t_1}(x)\epsilon_{t_2}(x) = \epsilon_{t_1+t_2}(x) \quad (t_1, t_2 \in \mathbb{M}, x \in \mathbb{Z})$$

azonosságot figyelembe véve adódik, hogy $(\mathbb{Z}, +)$ csoport duálisa izomorf az $(\mathbb{M}, \dot{+})$ csoporttal. (2.9) alkalmazásával, n-dimenziós esetben a $(\mathbb{Z}^n, +)$ karakter csoportját az $(\mathbb{M}^n, \dot{+})$ csoport azonosítja.

A (2.11) alatt bevezetett Fourier-transzformáció $\mathbb{G} = \mathbb{Z}$ esetre vonatkozó megfelelője az $\ell^1(\mathbb{Z})$-beli sorozatok

$$\widehat{f}(x) = \sum_{n\in\mathbb{Z}} f(n)exp(-2\pi inx) \quad (f \in \ell^1(\mathbb{Z}), x \in \mathbb{M}) \tag{3.11}$$

úgynevezett *Trigonometrikus Fourier-Sora* (TFS).

A (2.18) alatt bevezetett konvolúció TFS-re vontakozó alakja az alábbi alakot ölti:

$$(f * g)(m) := \sum_{n\in\mathbb{Z}} f(n)g(m-n) \quad (m \in \mathbb{G} = \mathbb{Z}, f, g \in \ell^1(\mathbb{Z})). \tag{3.12}$$

Itt jegyezzük meg, hogy minden – későbbiekben bemutatásra kerülő – líneáris, időinvariáns szűrés, tulajdonképpen egy rögzített függvénnyel vett konvolúció, azaz szemléletesen minden kimenő érték, a bemeneti értékek időtől független szűrő-együtthatókkal vett súlyzott összege. Ezt a fajta konvolúciót természetesen diszkrét jelek feldolgozásánál használjuk, számos bemenő jel ugyanis felfogható egy diszkrét sorozatnak, gondoljunk csak egy képre, ami véges sok pixel-értékből áll. Mint általánosságban, most is igaz a (2.19) formula, ami a konvolúció és a Fourier-transzformáció közötti kapcsolatot teremti meg.

Az (2.24) inverziós formula $\mathbb{G} = \mathbb{Z}$ csoportra vonatkozó alakja az

$$f(n) := \int_{\mathbb{M}} \widehat{f}(x)\epsilon_n(x)\, dx \quad (f \in \ell^1(\mathbb{Z})) \tag{3.13}$$

alakot ölti, ami tulajdonképpen az eredeti sorozat együtthatókból való előállítását jelenti.

3.4. Diszkrét Fourier-Sor

A periodikus sorozatok több szempontból is nagyon fontosak. Egyrészt, minden véges hosszúságú sorozat felfogható egy periodikus sorozat egyetlen periódusaként. Másrészt, a periodikus sorozatok úgynevezett *Diszkrét Fourier-Soros* reprezentációja képezi az összekötő láncszemet a teoretikus Fourier-integrálok és a numerikus számításra alkalmas diszkrét Fourier-transzformáltak között.

Ha a T_0 periódusidőt N egyenlő részre osztjuk, és az így kapott $T = T_0/N$ időt tekintjük egységnyinek, akkor bármelyik $t = nT \to n$ időpillanathoz tartozó $\widetilde{f}(nT) = \widetilde{f}(n)$ diszkrét helyettesítési érték a (3.9)-ből a $T\omega_0 = 2\pi/N$ helyettesítéssel

$$\widetilde{f}(n) = \sum_{m=-\infty}^{\infty} c_m e^{im\omega_0 nT} = \sum_{m=-\infty}^{\infty} c_m \epsilon_{mn}^N \quad n = 0, 1, \ldots, N-1$$

alakú lesz, ahol $\epsilon_n^N := \exp(in2\pi/N)$ a komplex egységvektor N-edik gyökét jelenti. A kapott sor m futóindexe szerinti összegzés csoportokra bontható a következő felbontással:

$$m = k + rN \quad k = 0, 1, \ldots, N-1 \text{ és } r \in \mathbb{Z}. \tag{3.14}$$

Az $\epsilon_N^N = 1$ és $\epsilon_{(k+rN)n}^N = \epsilon_{kn}^N$ azonosságok alkalmazásával a fenti sorfejtés tovább bontható, vagyis

$$\widetilde{f}(n) = \sum_{k=0}^{N-1} \sum_{r=-\infty}^{\infty} c_{k+rN} \epsilon_{(k+rN)n}^N = \sum_{k=0}^{N-1} \epsilon_{nk}^N \sum_{r=-\infty}^{\infty} c_{k+rN}. \tag{3.15}$$

Bevezetve a

$$\mathcal{F}\widetilde{f}(k) = N \sum_{r=-\infty}^{\infty} c_{k+rN} \quad k = 0, 1, \ldots, N-1 \tag{3.16}$$

úgynevezett *átlapolt együtthatókat*, akkor a (3.15) sorfejtés a

$$\widetilde{f}(n) = \sum_{k=0}^{N-1} \frac{1}{N} \mathcal{F}\widetilde{f}(k) \epsilon_{nk}^N, \quad n = 0, 1, \ldots, N-1 \tag{3.17}$$

véges periodikus összeg alakba írható, amiből a periodikus $\mathcal{F}\widetilde{f}(n)$ együtthatók meghatározhatók. Összefoglalva tehát, a (2.11) alatt bevezetett Fourier-transzformáció, periodikus $\widetilde{f}(n)$ időtartománybeli sorozatokra vonatkozó megfelelője a periodikus frekvencia-tartománybeli sorozatok

$$\widetilde{F}(k) = \mathcal{F}\widetilde{f}(n) = \sum_{n=0}^{N-1} \widetilde{f}(n) \epsilon_{-kn}^N, \quad k = 0, 1, \ldots, N-1 \tag{3.18}$$

úgynevezett *Diszkrét Fourier-Sora* (DFS).

A (2.18) alatt bevezettet konvolúció periodikus sorozatokra vonatkozó alakja a

$$(\widetilde{f} * \widetilde{g})(n) := \sum_{m=0}^{N-1} \widetilde{f}(m)\widetilde{g}(n-m) \tag{3.19}$$

formulával megadott, úgynevezett *periodikus konvolúció.* A most megadott konvolúció az eddigiektől annyiban különbözik, hogy mind az $\widetilde{f}(m)$ mind a $\widetilde{g}(n-m)$ N-re periodikus, így tehát a szorzatuk is az. Az összegzést azonban csak egyetlen periódusra kell elvégezni. Természetesen a (2.19) alatt bevezetett azonosság most is teljesül, méghozzá abban a formában, hogy két periodikus sorozat periodikus konvolúciójának diszkrét Fourier-soros transzformáltja a két sorozat DFS transzformáltjának a szorzatát adja.

A (2.24) alatt definiált inveziós formula, periodikus sorozatokra vonatkozó alakja a

$$\widetilde{f}(n) = \frac{1}{N} \sum_{k=0}^{N-1} \widetilde{F}(k)\epsilon_{kn}^{N}, \quad n = 0, 1, \ldots, N-1 \tag{3.20}$$

formában jelenik meg, mellyel az idő- és frekvencia-tartománybeli periodikus sorozatok között a dualitást teljessé tettük.

3.5. Diszkrét Fourier-Transzformáció

A véges hosszúságú sorozatok nagyon fontos esetében a DFS, illetve az inverz DFS Fourier-reprezentáció egyszerűsödik és az úgynevezett *Diszkrét Fourier-Transzformációba* és inverzébe megy át. Egy véges hosszúságú, N mintából álló $f(n)$ sorozat úgy hozható kapcsolatba a periodikus sorozatokat leíró DFS reprezentációval, hogy azt egy N pontos periódussal rendelkező periodikus $\widetilde{f}(n)$ jelsorozat egyetlen periódusának tekintjük. Az $f(n)$ sorozat periodikus ismétlésével generált $\widetilde{f}(n)$ (periodikus) sorozathoz viszont már tartozik egy egyértelmű $\widetilde{F}(k)$ DFS-együtthatós leírás, amely maga is N-re periodikus. Ha most ezekből az $\widetilde{F}(k)$ együtthatókból csak egyetlen periódust, mint véges hosszúságú N pontos $\widehat{f}(k)$ sorozatot tekintünk, akkor az előbbi eljárás fordított irányban is megismételhető.

A (2.3) alatt bevezetett karakterek vizsgálatával folytatva könnyen igazolható, hogy az $(\mathbb{M}_N, \dot{+})$ csoport karakterei az

$$\epsilon_n : \mathbb{M}_N \to \mathbb{T} \quad (n = 0, 1, \ldots, N-1)$$

alakú függvények és ezen csoport duálisa izomorf az egész számok

$$\mathbb{Z}_N := \{0, 1, \ldots, N-1\}$$

részhalmazának *mod* N vett csoportjával.

A (2.11) alatt bevezetett Fourier-transzformáció $\mathbb{G} = \mathbb{M}_N$ esetre vonatkozó megfelelője az $\ell^1(\mathbb{Z})$-beli diszkrét függvények

$$\widehat{f}(n) = \frac{1}{N} \sum_{t \in \mathbb{M}_N} f(t) exp(-2\pi int) \quad (f : \mathbb{M}_N \to \mathbb{C}, n \in \mathbb{Z}_N) \tag{3.21}$$

úgynevezett *Diszkrét Fourier-Transzformációja* (DFT). Ez a – gyakran találóan *véges Fourier-transzformációnak* is nevezett – mátrixkapcsolat N darab időtartományi komplex minta és N darab frekvencia-tartománybeli komplex spektrumvonal között teremt kölcsönösen egyértelmű összefüggést. A transzformáció tehát egyetlen speciális – (3.1) ábrán látható – mátrixszorzás segíségével megadja egy diszkrét idejű jelsorozat frekvenciaspektrumának mintáit. A transzformáló mátrix speciális alakjának köszönhetően a DFT-t algoritmikusan is egyetlen egységként, alapműveletként kezelik, holott természetesen egy bolnyolult rendszerről van szó.

$$\begin{bmatrix} \widehat{f}(0) \\ \widehat{f}(1) \\ \widehat{f}(2) \\ \vdots \\ \widehat{f}(N-1) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \epsilon^N & \epsilon^N_{-2} & \cdots & \epsilon^N_{-(N-1)} \\ 1 & \epsilon^N_{-2} & \epsilon^N_{-4} & \cdots & \epsilon^N_{-2(N-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \epsilon^N_{-(N-1)} & \epsilon^N_{-2(N-1)} & \cdots & \epsilon^N_{-(N-1)^2} \end{bmatrix} \cdot \begin{bmatrix} f(0) \\ f(1) \\ f(2) \\ \vdots \\ f(N-1) \end{bmatrix}$$

3.1. ábra. Az N-pontos DFT mátrixösszefüggése

A (2.18) alatt bevezetett konvolúció DFT-ra vontakozó alakja az alábbi formában írható fel:

$$(f * g)(x) := \frac{1}{N} \sum_{n \in \mathbb{M}_N} f(n) g(x-n) \quad (x \in \mathbb{G} = \mathbb{M}_N, f, g \in \ell^1(\mathbb{M}_N)). \tag{3.22}$$

A konvolóció elnevezésére gyakran használják a *diszkrét konvolúció* kifejezést, jelezve, hogy a súlyzott összeadás véges hosszúságú sorozatok felett történik.

A (2.24) inverziós formula $\mathbb{G} = \mathbb{M}_N$ csoportra vonatkozó alakja az

$$f(x) := \sum_{n \in \mathbb{Z}_N} \widehat{f}(n) \epsilon_n(x) \quad (f \in \ell^1(\mathbb{M}_N)) \tag{3.23}$$

alakot ölti, melynek mátrixösszefüggésére hasonló ábra rajzolható fel, a karakter komplex konjugálása nélkül.

A DFT, illetve IDFT jelölések helyett gyakran használják az FFT, illetve IFFT rövidítéseket is, amelyek – a *Fast Fourier-Transform* kifejezés rövidítéseként – a DFT-t és annak inverzét kölünös gyorsasággal előállító algoritmikus transzformációra utalnak. Tulajdonképpen ez a módszer (algoritmus) indította meg a modern jelfeldogozás térhódítását. Az FFT működését és annak megvalósítását a 4 fejezetben tárgyaljuk részletesebben.

3.6. A z-Transzformáció

A most következő szakaszban egy olyan transzformációt vezetünk be, amely – amellett, hogy a diszkrét Fourier-transzformáció általánosításának számít – matematikai módszerét illetően jól kezelhető, egyszerű és a diszkrét idejű jelfeldolgozás nélkülözhetetlen eszköze. *Diszkrét idejű rendszernek* egy olyan áramkörileg megvalósított, illetve algoritmikusan végrehajtott transzformációt nevezünk, amely a bemeneti $x(n)$ sorozatot egy kimeneti $y(n)$ sorozatba viszi át. A definíciót formálisan az $\mathcal{R}$ (Response) operátor segíségével írható fel a

$$y(n) = \mathcal{R}(x(n)) \tag{3.24}$$

formula segítségével. A további vizsgálataink során feltesszük, hogy a rendszer *líneáris* és *időinvariáns* [1]. Ismeretes, hogy bármely $x(n)$ sorozat összerakható az egységnyi impulzus késleltetett, illetve konstanssal szorzott értékeinek összegeként, azaz felírható egy

$$x(n) = \sum_{k=-\infty}^{\infty} x(k)\delta(n-k) \tag{3.25}$$

konvolúciós összeg alakjában, ahol $\delta(n) = 1$, ha $n = 0$ és azonosan 0 minden más n-re, az úgynevezett *egységimpulzus*. $h(n)$-nel jelölve az egységimpulzusra adott válasz-sorozatot, az időinvariancia a következő állítással ekvivalens: ha $h(n) = \mathcal{R}(\delta(n))$, akkor $h(n-k) = \mathcal{R}(\delta(n-k))$ is teljesül. Ennek és a (3.25) előállításnak a következményeképpen, a kimeneti $y(n)$ jel előállítható

$$y(n) = \mathcal{R}\left(\sum_{k=-\infty}^{\infty} x(k)\delta(n-k)\right) = \sum_{k=-\infty}^{\infty} x(k)\mathcal{R}(\delta(n-k)) = \sum_{k=-\infty}^{\infty} x(k)h(n-k), \tag{3.26}$$

azaz röviden $y(n) = x(n) * h(n)$, konvolúciós alakban.

A $h(n)$ impulzusválasz-sorozat meghatározása azonban sokszor nehézkes, ezért gyakran használják az úgynevezett *transzferfüggvényt* [1] a rendszer jellemzésére. A transzferfüggvény definíciószerűen egy olyan $H(z)$ függvény, amelyik egy komplex z változó függvényében egyértelműen megadja a rendszernek a z^n típusú gerjesztésre adott feleletét $z^n H(z)$ alakban. Behelyettesítve a fenti (3.26) egyenletbe az $x(n) = z^n$ sorozatot, láthatjuk, hogy az ilyen típusú sorozatok a rendszer sajátfüggvényeinek tekinthetők. Ugyanis ekkor a konvolúciós összeg a

$$y(n) = h(n) * x(n) = \sum_{k=-\infty}^{\infty} z^{n-k} h(k) = z^n \left(\sum_{k=-\infty}^{\infty} z^{-k} h(k) \right) = z^n H(z), \tag{3.27}$$

alakba írható, ahol a középső tagban a zárójelben lévő összeg:

$$H(z) = \sum_{k=-\infty}^{\infty} h(k) z^{-k}, \tag{3.28}$$

ami tulajdonképpen nem más, mint a bevezetni kívánt *z-transzformáció*.

Egy $x(n)$ diszkrét idejű sorozat $X(z)$ z-transzformáltját tehát a

$$X(z) = (\mathcal{Z}x)(n) = \sum_{n=-\infty}^{\infty} x(n) z^{-n} \tag{3.29}$$

sor segítségével definiáljuk, minden olyan z komplex értékre, amelyre ez a sor konvergens. Az előállítás tulajdonképpen egy *Laurent-sorfejtés* [12], amit *kétoldalú* z-transzformációnak is szoktak nevezni, megkülönböztetésül az $n = 0$-tól induló úgynevezett *egyoldalú* z-transzformációtól. Mivel kauzális jelek esetén a két előállítás ekvivalens, ezért a továbbiakban a jobban kezelhető kétoldalút használjuk. Egy rendszer *kauzális*, ha egy teszőleges $n = n_0$ időhöz tartozó kimeneti érték csak a korábbi, az $n \leq n_0$ időkhöz tartozó bemeneti minták értékétől függ. Ennek következményeként, egy a (3.26) szerinti konvolúciós összeggel jellemezhető líneáris, időinvariáns rendszer akkor és csak akkor kauzális, ha $h(n) = 0$, amikor $n < 0$.

A (3.29) sor konvergenciájának vizsgálatát leegyszerüsíti a Laurent-sorral való kapcsolat, ugyanis így lehetőség nyílik a komplex függvénytan eredményeinek felhasználására. Egy komplex $f(z)$ függvényt *analitikusnak* nevezünk egy z_0 pontban, ha az z_0 tetszőleges $|z - z_0| < \delta$ $(\delta > 0)$ környezetében differenciálható. Az analitikusság az $f(z) = u(x, y) + iv(x, y)$ felbontáshoz tartozó, úgynevezett *Cauchy-Riemann egyenletekből* [12] következik, ahol $z = x + iy$. Ha f a z_0 pontban nem analitikus, de minden környezetében található olyan pont, ahol $f(z)$ differenciálható, akkor a z_0 helyet a függvény *szinguláris* pontjának nevezzük.

A *Laurent-tétel* [12] szerint teszőleges a $d : r < |z - a| < R$ körgyűrűben egyértékű, analitikus $f(z)$ függvény – ebben a gyűrűben – előállítható a

$$f(z) = \sum_{n=-\infty}^{\infty} c_n (z - a)^n \tag{3.30}$$

Laurent-sorával, ahol a c_n konstansok a

$$c_n = \frac{1}{2\pi i} \oint_{\Gamma} f(\zeta)(\zeta - 0)^{-n-1}\, d\zeta, \quad n \in \mathbb{Z} \tag{3.31}$$

vonalintegrál formulával számolhatók, ahol Γ a D körgyűrűben fekvő tetszőleges, az óramutató járásával ellentétes irányítású görbe. Továbbá az R sugár növelésével, illetve a r sugár csökkentésével a D tartományt az $f(z)$ függvény legközelebbi külső, illetve belső szinguláris pontjáig növelhetjük, amely D' tartományon a sor (egyenletesen) konvergens és előállítja $f(z)$-t. Végül az $f(z)$ Laurent-sora a konvergencia tartományában egyértelmű, azonban az azonos középpontokhoz tartozó és szinguláris ponttal elválasztott konvergencia-tartományokban az $f(z)$ különböző Laurent-sorokkal állítható elő.

A komplex z változót az $e^{i\omega}$ egységkörön felvéve egyszerű kapcsolatot találunk egy $x(n)$ sorozat $X(z)$ és $\widehat{x}(n)$ z-, illetve Fourier-transzformáltja között:

$$X(e^{i\omega}) = \widehat{x}(n) = X(z)\big|_{z=e^{i\omega}} \ . \tag{3.32}$$

Altalánosabban, tekintsük a z komplex változó polárkoordinátás $z = e^{i\omega}$ alakját. Ekkor a z-transzformáció ezekhez a pontokhoz az $(x(n)r^{-n})$ sorozat Fourier-transzformáltját rendeli, ahonnan az $r = 1$ speciális esetben visszakapjuk az egységkörön vett (Fourier-)transzformáltat. Azonban a z-transzformált sorösszege akkor is konvergens lehet, amikor önmagában az $x(n)$ nem az, hiszen most a

$$\sum_{n=-\infty}^{\infty} |x(n)r^{-n}| < \infty$$

megkötésnek kell teljesülnie ahhoz, hogy a z-transzformált egyenletesen konvergens legyen. Ez a feltétel – egy adott $x(n)$ sorozat esetében – nyilvánvalóan csak az r, illetve a z bizonyos érték-tartományában fog teljesülni. Ha $X(z)$ konvergens a $r < |z| < R$ egyenlőtlenségekkel behatárolt körgyűrűben, akkor a z-nek ezt a halmazát *konvergenciatartománynak* nevezzük, melynek megadásával a transzformált már egyértelművé válik.

A Fourier-transzformációval való – most bevezett – szoros kapcsolat miatt magától értetődőnek látszik, hogy az ott bevezetett fogalmak a z-transzformációnál is érvényesek legyenek. Nos, a helyzet majdnem ez. A „majdnem" szó – mint látni fogjuk – a fentebb említett konvergenciatartományok miatt szerepel. Természetesen a z-transzformáció is egy líneáris operátor, amelynél a líneáris kombinációt alkotó sorozatok transzformáltjának konvergenciatartománya az egyes sorozatok konvergenciatartományainak a metszete.

A transzlációval való kapcsolat a következő formában érvényesül:

$$\mathcal{Z}(\tau_{-m}x(n)) = \sum_{n=-\infty}^{\infty} x(n-m)z^{-n} = \sum_{k=-\infty}^{\infty} x(k)z^{-(m+k)} = \tag{3.33}$$

$$= z^{-m}\sum_{k=-\infty}^{\infty} x(k)z^{-k} = z^{-m}X(z),$$

ahonnan a $z = e^{i\omega}$ helyettesítéssel visszakapjuk a (2.15) alatt bevezetett – Fourier-transzformációra vonatkozó – összefüggést. A konvergenciatartomány a transzláció hatására nem változik meg.

Legyen $\mathcal{Z}(x(n)) = X(z)$ a $D : r < |z| < R$ konvergenciatartományban. Ekkor az időtartományban a komplex w^{-n} sorozattal való szorzásnak a z-transzformált w-val való *dilatáltja* felel meg, nevezetesen:

$$\mathcal{Z}(x(n)w^{-n}) = \sum_{n=-\infty}^{\infty} x(n)w^{-n}z^{-n} = \sum_{n=-\infty}^{\infty} x(n)(wz)^{-n} = X(wz) = (\delta_w X)(z), \tag{3.34}$$

ahol a w egy tetszőleges komplex szám. A D konvergenciatartomány a w mennyiséggel átskálázódik: $D_w : r/|w| < |z| < R/|w|$.

Két sorozat *valós konvolúciójának* z-transzformáltjára a (2.19) összefüggéshez hasonló van érvényben, nevezetesen most a

$$\mathcal{Z}(x(n) * y(n)) = \mathcal{Z}(y(n) * x(n)) = X(z)Y(z) \tag{3.35}$$

formula teljesül, melynek konvergenciatartományára a líneraritásnál elmondottak teljesülnek.

A $\mathcal{Z}^{-1}$ inverz z-transzformáció egy $X(z)$ z-transzformálthoz egy $x(n)$ sorozatot rendel az alábbi módon:

$$x(n) = \mathcal{Z}^{-1}(X(z)) = \frac{1}{2\pi i}\oint_C X(z)z^{n-1}\,dz, \quad (n \in \mathbb{Z}), \tag{3.36}$$

ahol C egy olyan kontúrgörbe, amely az $X(z)$ konvergenciatartományában, az óramutató járásával ellentétes irányban haladva körülfogja $X(z)$ szingularitásait, valamint az origót. A fenti vonalintegrál – a komplex függvénytanban jól ismert – *Reziduum-tétel* [12] segítségével könnyen meghatározható:

$$x(n) = \frac{1}{2\pi i}\oint_C X(z)z^{n-1}\,dz = \sum_{j=1}^{M} Res_{z=z_j}(X(z)z^{n-1}) = \mathcal{Z}^{-1}(X(z)), \quad (3.37)$$

ahol z_j-k ($j = 1, \ldots, M$) a C görbe belsejéban lévő pólusok.

Arról már volt szó, hogy két *valós* sorozat konvolúcióját a z-transzformáció a transzformáltak szorzatába viszi át. Most vizsgáljuk meg két teszőleges $x(n)$ és $y(n)$ sorozat szorzatának z-transzformáltját. Ehhez legyenek az egyes sorozatokhoz tartozó transzformáltak $X(z)$, illetve $Y(z)$, melyek a D_x, illetve D_y tartományban konvergensek. Ekkor a sorozatok szorzataként definiált $w(n)$ sorozat $W(z)$ z-transzformáltját az alábbi *komplex konvolúció* segítségével számoljuk:

$$W(z) = \mathcal{Z}(w(n)) = \mathcal{Z}(x(n)y(n)) = \frac{1}{2\pi i}\oint_C Y(v)X(\frac{z}{v})v^{-1}\,dv, \qquad (3.38)$$

ahol C az óramutató járásával ellentétes irányítású, D_x és D_y tartományok metszetének konvergenciatartományában haladó zárt görbe. A most bevezetett (3.38) formulából közvetlenül levezethető a Parseval-féle formula:

$$\sum_{n=-\infty}^{\infty} x(n)\overline{y}(n) = \frac{1}{2\pi i}\oint_C X(v)\overline{Y}(\frac{1}{\overline{v}})v^{-1}\,dv, \qquad (3.39)$$

ahol $\overline{v}$ a v komplex konjugáltját jelöli.

A korábbiakban láthattuk, hogy egy periodikus $\widetilde{x}(n)$ sorozatot az ugyancsak periodikus $\widetilde{X}(k)$ diszkrét Fourier-soros együtthatók segítségével írhatunk le, továbbá ez a kapcsolat a (3.18) és (3.20) szerint kölcsönös és egyértelmű. Nyílvánvaló, hogy a periodikus $\widetilde{x}(n)$ sorozatnak nem létezik z-transzformáltja, hiszen a periodicitás miatt az előállító sor sehol sem fog konvergálni. Helyette emeljük ki az N-re periodikus sorozat egy periódusát és jelöljük $x(n)$-nel, azaz $x(n) = \widetilde{x}(n)$, ha $0 \leq n \leq N-1$, és $x(n) = 0$ egyébként. Ez által a (3.29) alatt bevezetett z-transzformált a

$$X(z) = \sum_{n=-\infty}^{\infty} x(n)z^{-n} = \sum_{n=0}^{N-1} \widetilde{x}(n)z^{-n} \qquad (3.40)$$

véges összegre egyszerűsödik. Ezt összevetve a (3.18) alatti, $\widetilde{X}(k)$ DFS együtthatók meghatározására szolgáló egyenlettel, megfigyelhető, hogy $X(z)$ és $\widetilde{X}(k)$ kapcsolatára az

$$\widetilde{X}(k) = X(z)\Big|_{z=e^{i\frac{2\pi k}{N}}} = X(e^{i\frac{2\pi k}{N}}) = X(\epsilon_k^N) \qquad (3.41)$$

összefüggés írható fel. Látható tehát, hogy az $\widetilde{X}(k)$ DFS együtthatók az $X(z)$ z-transzformáltnak a komplex egységkörön felvett mintáival egyenlők, amely N darab minta a $z = 1$ pontból kiindulva egymástól $2\pi/N$ radiánra helyezkedik el.

Most vizsgáljuk meg a kapcsolatot a másik oldalról is, azaz induljunk ki egy olyan nemperiodikus, általában véges hosszúságú $x(n)$ sorozatból, amelynek létezik a (3.29) szerinti $X(z)$ z-transzformáltja, amely egyértelműen adott. Ha most az $X(z)$ egységkörös mintáit a (3.41) alatti összefüggéshez hasonlóan az

$$\widetilde{X}(k) = X(z)\Big|_{z=e^{i\frac{2\pi k}{N}}} = \sum_{m=-\infty}^{\infty} x(m)e^{-i\frac{2\pi}{N}km} = \sum_{m=-\infty}^{\infty} x(m)\epsilon_{-km}^{N} \tag{3.42}$$

formában DFS együtthatóknak tekintjük, akkor az inverz DFS transzformáció egy periodikus

$$\widetilde{x}(n) = \frac{1}{N}\sum_{k=0}^{N-1} \widetilde{X}(k)e^{i\frac{2\pi}{N}kn} = \frac{1}{N}\sum_{k=0}^{N-1} \widetilde{X}(k)\epsilon_{kn}^{N} \tag{3.43}$$

sorozatot rendel hozzá. A (3.42) alatti összefüggést behelyettesítve a most levezetett eredménybe a

$$\widetilde{x}(n) = \frac{1}{N}\sum_{k=0}^{N-1}\left(\sum_{m=-\infty}^{\infty} x(m)\epsilon_{-km}^{N}\right)\epsilon_{kn}^{N} = \sum_{m=-\infty}^{\infty} x(m)\left(\frac{1}{N}\sum_{k=0}^{N-1}\epsilon_{k(n-m)}^{N}\right) \tag{3.44}$$

formulához jutunk, ahol a zárójelben lévő kifejezés értéke éppen 1, ha $m = n + rN$ és zérus, minden más m értékre. Ez utóbbi megjegyzés figyelembevételével végeredményül az

$$\widetilde{x}(n) = \sum_{r=-\infty}^{\infty} x(n + rN) \tag{3.45}$$

egyenlőséghez jutunk, azaz a z-transzformált mintavételezésével és az azt követő inverz DFS transzformációval előállított periodikus $\widetilde{x}(n)$ sorozat a kiinduló $x(n)$ sorozat N-pontos ismétlésével állítható elő. Azonban az előállítás csak akkor lehet egyértelmű, ha az $x(n)$ sorozat hosszúsága eredetileg kisebb volt, mint N. Ellenkező esetben ugyanis, $\widetilde{x}(n)$ előállítása közben a sorozatok *átlapolódnak* és a z-transzformált mintavételezéséből kapott $\widetilde{x}(n)$ sorozatból az $x(n)$ nem nyerhető vissza. A probléma természetesen megoldható N növelésével.

Most a z-transzformáltat az egységkörű mintáiból előállítjuk zárt alakban. Kindulva az N pont hosszúságú $x(n)$ sorozat $X(z)$ z-transzformáltjából:

$$X(z) = \sum_{n=0}^{N-1} x(n) z^{-n}, \tag{3.46}$$

helyettesítsünk az $x(n)$ helyébe az $\widetilde{x}(n)$ sorozat (3.43) szerinti inverz DFS kifejtését (ez megtehető, hiszen $x(n) = \widetilde{x}(n)$, ha $0 \leq n \leq N-1$). Ekkor kapjuk, hogy

$$X(z) = \sum_{n=0}^{N-1} \left(\frac{1}{N} \sum_{k=0}^{N-1} \widetilde{X}(k) \epsilon_{kn}^{N} \right) z^{-n} = \frac{1}{N} \sum_{k=0}^{N-1} \widetilde{X}(k) \left(\sum_{n=0}^{N-1} (\epsilon_k^N z^{-1})^n \right). \tag{3.47}$$

A zárójelben lévő kifejezés megadható zárt alakban a mértani sor részletösszegére vonatkozó formula segítségével, továbbá a $\epsilon_{kN}^{N} = 1$ miatt az $X(z)$ z-transzformált a következő alakot ölti:

$$X(z) = \frac{1}{N} \sum_{k=0}^{N-1} \widetilde{X}(k) \frac{1 - z^{-N}}{1 - \epsilon_k^N z^{-1}} = \frac{1 - z^{-N}}{N} \sum_{k=0}^{N-1} \frac{\widetilde{X}(k)}{1 - \epsilon_k^N z^{-1}} \quad . \tag{3.48}$$

Azonban a mintavételezést általában egy $r < 1$ sugarú körön szokták elvégezni, ezért a (3.48) egy általánosabb alakja:

$$X(z) = (1 - r^N z^{-N}) \frac{1}{N} \sum_{k=0}^{N-1} \frac{\widetilde{X}(r\epsilon_k^N)}{1 - r\epsilon_k^N z^{-1}} \quad . \tag{3.49}$$

Ez utóbbi formula kiszámítása a $\widetilde{X}(r\epsilon_k^N)$ minták meghatározása miatt nehézkes, azonban, ha az r elég közel van az egységhez, akkor számolhatunk az eredeti $\widetilde{X}(\epsilon_k^N)$ mintákkal is.

A most levezetett zárt alak segítségével a Fourier-transzformáltat is megadhatjuk ilyen formában a $z = e^{i\omega}$ helyettesítéssel. Ekkor ugyanis a (3.48) formula a

$$X(e^{i\omega}) = \sum_{k=0}^{N-1} \widetilde{X}(k) \Phi(\omega - \frac{2\pi}{N} k) \tag{3.50}$$

alakba hozható. Kifejtve a (3.48) alatti összeg k-adik tagját:

$$\frac{1}{N} \frac{1 - z^{-N}}{1 - \epsilon_k^N z^{-1}} = \left. \frac{z^{-\frac{N}{2}} \left(z^{\frac{N}{2}} - z^{-\frac{N}{2}} \right)}{N \left(z^{\frac{1}{2}} \epsilon_{-\frac{k}{2}}^{N} - z^{-\frac{1}{2}} \epsilon_{\frac{k}{2}}^{N} \right) \epsilon_{\frac{k}{2}}^{N} z^{-\frac{1}{2}}} \right|_{z=e^{i\omega}} =$$

$$= e^{-i\omega \frac{N-1}{2}} \frac{\sin(N\omega/2)}{N e^{i\pi k/N} \sin(\omega/2 - \pi k/N)} \quad , \tag{3.51}$$

majd a $k = 0$ esethez tartozó számolgatások után $\Phi(\omega)$ – interpoláló függvény – explicit kifejezhető a

$$\Phi(\omega) = e^{-i\omega\frac{N-1}{2}} \frac{\sin(N\omega/2)}{N\sin(\omega/2)} \tag{3.52}$$

formula segítségével. Összefoglalva a kapott eredményeket a Fourier-transzformált az alábbi zárt alakra hozható:

$$X(e^{i\omega}) = \sum_{k=0}^{N-1} \widetilde{X}(k) e^{-i\omega\frac{N-1}{2}} \frac{\sin(N\omega/2)}{Ne^{i\pi k/N}\sin(\omega/2 - \pi k/N)} \quad . \tag{3.53}$$

4. fejezet

Algoritmusok

4.1. Gyors Fourier-Transzformáció

A 3.5. szakaszban bevezetett DFT azért válhatott a diszkrét idejű jelfeldolgozás és szűrés meghatározó elemévé, mert léteznek olyan algoritmusok, illetve számítástechnikai megoldások, amelyekkel a DFT operációinak hatalmas műveletigényei igen nagy mértékben lecsökkenthetők, ezzel sokkal hatékonyabbá és gyorsabbá téve a transzformációt. Egy N-pontos DFT közvetlen kiszámítása során bármelyik $X(k)$ DFT együttható kiszámításához N komplex szorzást és $N-1$ komplex összeadást kell elvégezni. Ez azt jelenti, hogy az N darab együttható meghatározásához összesen N^2 komplex szorzásra és $(N-1)N$ komplex összeadásra van szükség, tehát a transzformáció műveletigénye $\mathcal{O}(N^2)$. Természetesen a komplex műveleteket a jól ismert

$$\epsilon_k(x) = e^{i2\pi kx} = \cos(2\pi kx) + i\sin(2\pi kx), \quad (k \in \mathbb{Z}, x \in \mathbb{R}) \tag{4.1}$$

Euler-formula segítségével felbonthatjuk valós operációkra a következő képpen:

$$\begin{aligned} X(k) &= \sum_{n=0}^{N-1} x(n)\epsilon_{-nk}^{N} = \sum_{n=0}^{N-1} x(n)\left(\cos\left(\frac{2\pi nk}{N}\right) - i\sin\left(\frac{2\pi nk}{N}\right)\right) = \\ &= \sum_{n=0}^{N-1}\left[\left(\Re(x(n))\cos\left(\frac{2\pi nk}{N}\right) + \Im(x(n))\sin\left(\frac{2\pi nk}{N}\right)\right) - \right. \\ &\quad \left. - i\cdot\left(\Re(x(n))\sin\left(\frac{2\pi nk}{N}\right) - \Im(x(n))\cos\left(\frac{2\pi nk}{N}\right)\right)\right], \end{aligned} \tag{4.2}$$

ahol $\Re(x)$, illetve $\Im(x)$ a komplex x szám valós, illetve képzetes részét jelölik. Az így kialakított valós operációk műveletigénye $4N^2$ *valós szorzás* és

$N(4N-2)$ *valós összeadás*. Mindehhez hozájön még a komplex $x(n)$ sorozat $2N$ valós mintájának, valamint az ϵ^N_{-nk} tényezőknek megfelelő színusz- és koszinusz-együtthatók sorozatának tárolása. Ezek után nem nehéz belátni, hogy az N-pontos DFT direkt módon történő kiszámítása N^2-tel arányosan nő. Ezzel szemben a – most bemutatásra kerülő – *gyors Fourier-transzformációk (FFT)*, egy ugyanilyen N elemű bemenet DFT kiszámításához, csak $N \log_2 N$ komplex műveletet végeznek. Ez a $\log_2 N/N$ arányú műveletigény-csökkenés méltán tekinthető radikálisnak, hiszen már egy $N = 1024$ elemű bemeneti minta transzformációjának kiszámítása is kb. 100-szoros javulást eredményez.

A következőkben egy példán kersztül szemléltetjük a DFT redundanciáját. Legyen a bemenetünk egy $N = 8$ pont hosszúságú, zérusokkal 16 eleműre kiegészített valós sorozat. A sorozat valóssága miatt a (4.2) alatt levezetett összefüggés az alábbi

$$X(k) = \sum_{n=0}^{15} x(n) \cos\left(\frac{nk\pi}{8}\right) - \; i \cdot \sum_{n=0}^{15} x(n) \sin\left(\frac{nk\pi}{8}\right) \qquad (4.3)$$

egyszerűbb alakba írható. Ismeretes, hogy a DFT-re érvényes az úgynevezett *komplex konjugációs tétel* [1], azaz fennáll a

$$X(N-k) = \overline{X}(k), \quad (k = 0, 1, 2, \ldots, \frac{N}{2}) \qquad (4.4)$$

egyenlőség, ahol $\overline{z}$ a z szám komplex konjugáltját jelöli és N páros szám. A fenti (4.3) alatti egyenlőség $k = 0, 1, \ldots, 8$ értékek melletti, pl. $\cos(nk\pi/8)$ számokat tartalmazó, 16 oszlopos C mátrix sorainak száma, az iménti (4.4) tétel következményeként, 9-re redukálódik. Ez által a szögfüggvények értékeit tartalmazó C mátrix a 4.1. ábrán látható alakba írható, ahol az a, b és c speciális konstansok érétkei rendre $\cos(\pi/8)$, $\cos(2\pi/8)$ és $\cos(3\pi/8)$. E mátrix segítségével a DFT a következő mátrix-vektor szorzatként állítható elő:

$$X = Cx,$$

ahol x a 16 elemű mintát tartalmazó bemenő jel, X pedig a DFT együtthaókat tartalmazó 9 elemű tömb. A mátrixszorzás úgy interpretálható, hogy az $x(n)$ bemenő jelet 9 különböző koszinusz-generátor jellel összehasonlítjuk és $k = 0, 1, \ldots, 8$ értékekre megvizsgáljuk, hogy a k-adik vizsgáló koszinusz-jel „mennyire hasonlít” az $x(n)$ sorozatra. A hasonlóság mértékén most egy diszkrét skaláris-szorzatot értünk, nevezetesen a C mátrix k-adik sorát jobbról megszorozzuk az $x(n)$ mintákból álló x vektorral, mely szorzatnak az eredménye éppen az $X(k)$ DFT frekvenciaminta valós része lesz. Ha most

$$\begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & a & b & c & 0 & -c & -b & -a & -1 & -a & -b & -c & 0 & c & b & a \\
1 & b & 0 & -b & -1 & -b & 0 & b & 1 & b & 0 & -b & -1 & -b & 0 & b \\
1 & c & -b & -a & 0 & a & b & -c & -1 & -c & b & a & 0 & -a & -b & c \\
1 & 0 & -1 & 0 & 1 & 0 & -1 & 0 & 1 & 0 & -1 & 0 & 1 & 0 & -1 & 0 \\
1 & -c & -b & a & 0 & -a & b & c & -1 & c & b & -a & 0 & a & -b & -c \\
1 & -b & 0 & b & -1 & b & 0 & -b & 1 & -b & 0 & b & -1 & b & 0 & -b \\
1 & -a & b & -c & 0 & c & -b & a & -1 & a & -b & c & 0 & -c & b & -a \\
1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1
\end{bmatrix}$$

4.1. ábra. A $\cos(nk\pi/8)$ értékeket tartalmazó C mátrix

az eddigi 16 pontos DFT helyett csak egy 8 elemes DFT-t számolunk, akkor a szorzásban a C mátrixnak csak minden második oszlopa és első nyolc sora vesz részt. Ez által a változatlan $x(n)$ infomációtartalom mellett kétszeres fokszámú DFT meghatározást tettünk lehetővé, hiszen az $x(n)$ minta második nyolc elemét zérusnak definiáltuk. Természetesen a DFT együtthatók képzetes – színuszos – részeire is hasonló megállapítás tehető, egy kevésbé eltérő másik mátrix felírásával. A sok 1, -1 és 0 értéket tartalmazó mátrix miatt a szorzások száma érezhetően csökkenni fog, sőt a három speciális konstanssal való szorzás is értelemszerűen csoportosítható.

A fent leírtak alapján tényként kezelhető, hogy a DFT direkt kiszámításánál a trigonometrikus szögfüggvények tulajdonságai miatt, mind az n, mind a k indexek vonatkozásában nagyfokú redundanciával dolgozunk. A transzfomációs mátrixokban a sorok, illetve oszlopok szerinti ismétlődés nyilvánvalóan N növelésével arányosan nőni fog. Ennek hatékony kiaknázásán alapul az összes FFT algoritmus alapgondolata, mely szerint az *N-pontos DFT-t kisebb pontszámú DFT-k segítségével kell elvégezni.* Az FFT tehát lebontások sorozataként valósítható meg, mely lebontási sorozatot most a legegyszerűbb *R2 DIT1* (Radix 2 Decimation In Time) [1] FFT algoritmuson mutatunk be. A Radix 2 jelző arra utal, hogy a bemeneti minta hossza kettő hatvány, azaz $N = 2^L$, $(L \in \mathbb{Z})$ alakba írható. Az In Time – tehát a DIT – elnevezés pedig azt jelöli, hogy az összekapcsolási műveletek a bementi időtartományi minták átrendezésével valósíthatók meg (a későbbiekben szó lesz még a frekvenciatartomány átrendezésével (DIF) megvalósítható algoritmusokról is).

Természetesnek tűnik az a gondolat, hogy tetszőleges páros N-hez tartozó DFT két olyan $(N/2)$-pontos DFT eredményének kombinálásával kapjunk meg, amelyek már együttesen is csak $2(N/2)^2 = N^2/2$ műveletet igényelnek. Mivel N kettő hatvány, ezért ez a fajta felbontás megismételehető, egészen addig, amíg csupa 2-pontos DFT-k kiszámításához nem jutunk. A szakirodalomban ezeket – a megvalósító hatásgráf alakjáról – *pillangóknak* (angolul *butterfly*) nevezik. Egy ilyen 2-pontos DFT szemléltetéséhez a két komplex bemenetet jelöljük x_1-gyel és x_2-vel, a komplex kimeneteket pedig X_1-gyel és X_2-vel. Ekkor a DFT a bemenetekhez tartozó kimeneteket a ϵ_{-k}^N együtthatók segítségével a következő egyenletek szerint állítja elő:

$$X_1 = x_1 + \epsilon_{-k}^N \cdot x_2, \tag{4.5}$$
$$X_2 = x_1 + \epsilon_{-k+N/2}^N \cdot x_2 = x_1 - \epsilon_{-k}^N \cdot x_2. \tag{4.6}$$

Észrevehető, hogy a pillangó csak egyetlen szorzást igényel, ugyanis az $(\epsilon_{-k}^N \cdot x_2)$ mennyiséget csak egyetlen egyszer kell kiszámolni. Mivel $N = 2^L$, ezért a kétpontos DFT pillangókig $L = \log_2 N$ darab egymás utáni felbontást kell alkalmazni. Minthogy mindegyik felbontásnál továbbra is N pontot – azaz (N/2) kétpontos pillangót – kell kiszámolnunk, ezért összesen $(N \log_2 N)/2$ pillangó – komplex szorzás – végrehajtására van szükség.

Természetesen az FFT algoritmust „helyesen" kell felépíteni, abban az értelemben, hogy az egyes lebontásoknál a felezett DFT-k jó bemeneti mintákat kapjanak és a kimeneti mintáikat a megfelelő helyekre juttassák el. Az N elemű $x(n)$ komplex bemeneti mintát bontsuk fel két $N/2$ hosszú $x_1(n)$ és $x_2(n)$ mintákra a következő képpen:

$$x_1(n) := x(2n) \text{ és } x_2(n) := x(2n+1) \quad n = 0, 1, \ldots, (N/2) - 1\ . \tag{4.7}$$

Alkalmazva a

$$\epsilon_{-2}^N = (e^{-i(2\pi/N)})^2 = e^{-i2\pi/(N/2))} = \epsilon_{-1}^{N/2} \tag{4.8}$$

azonosságot, valamint a (3.21) alatt bemutatott DFT definíciójának egy ekvivalens átfogalmazását, a

$$X(k) = \sum_{n=0}^{N-1} x(n)\epsilon_{-nk}^N = \sum_{n=0}^{(N/2)-1} x(2n)\epsilon_{-2nk}^N + \sum_{n=0}^{(N/2)-1} x(2n+1)\epsilon_{-(2n+1)k}^N =$$
$$= \sum_{n=0}^{(N/2)-1} x_1(n)\epsilon_{-nk}^{N/2} + \epsilon_{-k}^N \sum_{n=0}^{(N/2)-1} x_2(n)\epsilon_{-nk}^{N/2} = X_1(k) + \epsilon_{-k}^N X_2(k),$$
$$k = 0, 1, \ldots, (N/2) - 1 \tag{4.9}$$

formulához jutunk, ahol $X_1(k)$, illetve $X_2(k)$ már az $N/2$ pontos, felezett együtthatók. Látható, hogy eddig csak a $0 \leq k \leq (N/2) - 1$ esettel foglalkoztunk. Az $(N/2) - 1$-nél nagyobb k értékekre való kiterjesztéshez vegyük figyelembe az $X_1(k)$ és $X_2(k)$ sorozatok $N/2$ szerinti periodicitását és helyettesítsünk a (4.9) egyenletbe k helyére $(k - N/2)$ értéket:

$$X(k) = X_1(k - N/2) + \epsilon^N_{-k} X_2(k - N/2), \quad N/2 \leq k \leq N - 1 \tag{4.10}$$

azaz indexcserével:

$$X(k - N/2) = X_1(k) + \epsilon^N_{-k+N/2} X_2(k), \quad 0 \leq k \leq (N/2) - 1, \tag{4.11}$$

ahol az $\epsilon^N_{-k+N/2}$ formula a következő alakra hozható:

$$\epsilon^N_{-k+N/2} = \epsilon^N_{-k} \cdot \epsilon^N_{N/2} = -\epsilon^N_{-k} \ .$$

Ezek után – folytatva a következő fokozattal – definiáljuk $x_{11}(n)$, $x_{21}(n)$, $x_{12}(n)$ és $x_{22}(n)$ $(N/4)$ hosszúságú páros, illetve páratlan sorozatokat a következő képpen:

$$x_{11}(n) := x_1(2n), \quad x_{21}(n) := x_2(2n),$$

$$x_{12}(n) := x_1(2n + 1), \quad x_{22}(n) := x_2(2n + 1),$$
$$n = 0, 1, \ldots, (4/N) - 1 \ . \tag{4.12}$$

Ekkor, az eddigiek alapján a következő fokozatra igaz, hogy

$$X_1(k) = \begin{cases} X_{11}(k) + \epsilon^{N/2}_{-k} \cdot X_{12}(k), & k = 0, 1, 2, \ldots, \frac{N}{4} - 1, \\ X_{11}(k - \frac{N}{4}) + \epsilon^{N/2}_{-k} \cdot X_{12}(k - \frac{N}{4}), & \frac{N}{4} \leq k \leq \frac{N}{2} - 1, \end{cases}$$

illetve

$$X_2(k) = \begin{cases} X_{21}(k) + \epsilon^{N/2}_{-k} \cdot X_{22}(k), & k = 0, 1, 2, \ldots, \frac{N}{4} - 1, \\ X_{21}(k - \frac{N}{4}) + \epsilon^{N/2}_{-k} \cdot X_{22}(k - \frac{N}{4}), & \frac{N}{4} \leq k \leq \frac{N}{2} - 1 \ . \end{cases}$$

A fent leírt lebontási lépéseket addig folytatjuk, amíg csupa 2-pontos DFT együtthatókig el nem jutunk. Mivel $N = 2^L$, ezért L fokozaton keresztül ismételjük meg lépéseinket.

Az eddigiekben leírtuk, miként lehet a DFT együtthatók kiszámítását meggyorsítani, most ejtsünk szót az algoritmus tárhelyigényéről. Mint láttuk az algoritmus L lépésének mindegyikében, N darab komplex számból kiindulva, $(N/2)$ pillangó-műveletet végez el és N darab komplex számot eredményez. Minthogy az egyik fokozat pillangó-műveletének kimenete a másik (következő) fokozat pillangó-műveletének a bemenete, ésszerűnek látszik a *helyben számolás* követelményének megkötése. Mivel a bemenő sorozatra már nincs szükségünk a további számolásokban, ezért azt az első fokozat kimeneti mintáival felülírhatjuk. Az így kapott sorozat szolgáltatja a második fokozat bemeneti mintáját. A következő $(N/2)$ pillangó-művelet után a kimenettel ismét felülírjuk a bemenetet. Ezeket a lépéseket természetesen L fokozaton keresztül megismételjük, így az adott tárhelyen az algoritmus leállása után éppen a kívánt DFT együtthatók lesznek. A műveletek elvégzése után a sorrend kimeneten más lesz, mint a bemeneten, de mivel az algoritmus a futása során egyetlen fokozatban sem ír egy rekeszbe két elemet, ezért a helyben számolás megvalósítható, tehát csak N komplex szám tárolására van szükség. Azonban a helyben számolásnak is van ára, nevezetesen a bemeneti minta elemeinek sorrendjét át kell rendezni az úgynevezett *fordított bitképű* (bit-reversed) sorrendbe, a következő (4.13) transzformációval [3]. Legyen

$$x = \sum_{n=0}^{\infty} \frac{x_n}{2^{n+1}}, \quad m = \sum_{n=0}^{\infty} m_n 2^n, \quad (x_n, m_n \in \{0,1\})$$

az $x \in [0,1)$ és az $m \in \mathbb{N}$ számok diadikus – vagyis bináris – előállítása. A

$$\begin{aligned} \pi_N(x) &:= \frac{x_{N-1}}{2} + \frac{x_{N-2}}{2^2} + \cdots + \frac{x_0}{2^N} + \sum_{n=N}^{\infty} \frac{x_n}{2^{n+1}}, \\ \dot{\pi}_N(m) &:= m_{N-1} + m_{N-2}2 + \cdots + m_0 2^{N-1} + \sum_{n=N}^{\infty} m_n 2^n \end{aligned} \tag{4.13}$$

utasítással értelmezett $\pi_N : [0,1) \to [0,1)$ és $\dot{\pi}_N : \mathbb{N} \to \mathbb{N}$ leképezéseket *bitfordító transzformációnak nevezzük.* Szemléltetéshez példaként írjuk fel az n index értékét az (n_2, n_1, n_0) számhármassal (ahol $n_i \in \{0,1\}$, $(i \in \{0,1,2\})$) bináris formában. Ekkor az eredeti soros $x(n) = x(n_2 \cdot 2^2 + n_1 \cdot 2 + n_0)$ mintákból az $x(\dot{\pi}_3(n)) = x(n_0 \cdot 2^2 + n_1 \cdot 2 + n_2)$ bitfordított mintákat kell sorrendben kiválasztani és szorosan az FFT bementéhez kapcsolni.

A fejezet további részében a most bemutatott algoritmust általánosítjuk és bemutatunk még további algoritmus családokat.

4.2. FFT algoritmusok algebrai formája

Az általánosítás bevezetéséhez egy – az előzőekben megismert – hasonló lebontást használunk. Láthattuk, hogy egy N-pontos $x(n)$ sorozat $X(k)$ DFT együtthatóit egy $L = \log_2 N$ fokozatú FFT segítségével számolhatjuk ki. Az általánosabb forma felírásához az n és k számokat írjuk fel bínáris alakban:

$$n = \sum_{j=0}^{L-1} n_j 2^j = \sum_{j=0}^{L-1} n_{L-1-j} 2^{L-1-j} = n_{L-1} 2^{L-1} + \dots n_1 2^1 + n_0 2^0,$$

$$n_j \in \{0,1\}, j \in \{0,1,\dots,L-1\}, \tag{4.14}$$

illetve

$$k = \sum_{j=0}^{L-1} k_j 2^j = \sum_{j=0}^{L-1} k_{L-1-j} 2^{L-1-j} = k_{L-1} 2^{L-1} + \dots k_1 2^1 + k_0 2^0,$$

$$k_j \in \{0,1\}, j \in \{0,1,\dots,L-1\}. \tag{4.15}$$

Ekkor az $x(n)$ sorozatot felírhatjuk annak bínáris jegyeivel:

$$x(n) = x(n_{L-1} 2^{L-1} + \dots n_1 2 + n_0) =: x_0(n_{L-1}, n_{L-2}, \dots, n_1, n_0), \tag{4.16}$$

melynek segítségével a (3.21) DFT reprezentálható többszörös összegként:

$$X(k) = \sum_{n=0}^{N-1} x(n) \epsilon_{-kn}^{N} = \tag{4.17}$$

$$= \sum_{n_0=0}^{1} \sum_{n_1=0}^{1} \cdots \sum_{n_{L-1}=0}^{1} x_0(n_{L-1}, n_{L-2}, \dots, n_1, n_0) \epsilon^{N}_{-k(n_{L-1}2^{L-1}+\dots n_1 2 + n_0)} .$$

Az $N = 8$ esethez tartozó példán keresztül, most szemléltetjük, hogy miként lehet az összeget n bínáris jegyeitől (n_i) függetlenné tenni. Tehát erre a speciális esetre a (4.17) sor a

$$\begin{aligned} X(k) &= \sum_{n_0=0}^{1} \sum_{n_1=0}^{1} \sum_{n_2=0}^{1} x_0(n_2, n_1, n_0) \epsilon^8_{-4kn_2} \epsilon^8_{-2kn_1} \epsilon^8_{-kn_0} = \\ &= x_0(0,0,0) + x_0(0,0,1)\epsilon^8_{-k} + x_0(0,1,0)\epsilon^8_{-2k} + x_0(0,1,1)\epsilon^8_{-3k} + \\ &+ x_0(1,0,0)\epsilon^8_{-4k} + x_0(1,0,1)\epsilon^8_{-5k} + x_0(1,1,0)\epsilon^8_{-6k} + x_0(1,1,1)\epsilon^8_{-7k} \end{aligned}$$

konkrét formát ölti, ahol nyilvánvalóan x_0 elemei természetes sorrendben vannak indexelve. A fenti (4.17) összegzés legbelső tagját fejtsük ki a k-ra vonatkozó (4.15) bínáris felírás segítségével:

$$\sum_{n_2} x_0(n_2, n_1, n_0) \epsilon^8_{-4n_2(4k_2+2k_1+k_0)} = \sum_{n_2} x_0(n_2, n_1, n_0)(-1)^{k_0 n_2}, \tag{4.18}$$

ahol kihasználtuk, hogy $\epsilon_{-4}^{8} = -1$ és $(-1)^{n_2(4k_2+2k_1)} = 1$ tetszőleges n_2 mellett. Elvégezve az így kapott n_2 szerinti összegzést, észrevehető, hogy egy olyan $x_1(n)$ sorozathoz jutunk, amely már n_2 helyett k_0 értékétől függ, ami annyit jelent, hogy n_2 helyébe k_0 írható. Ezzel tulajdonképpen egy kétpontos DFT műveletet jelöltünk ki, ugyanis rögzített n_0 és n_1 értékek mellet az $x_1(k_0, n_1, n_0)$ kételemű sorozat, éppen az $x_0(n_2, n_1, n_0)$ kétpontos sorozat DFT együtthatóinak felel meg. Elvégezve tehát az n_2 szerinti összegzést a

$$x_1(k_0, n_1, n_0) = \sum_{n_2=0}^{1} x_0(n_2, n_1, n_0)(-1)^{k_0 n_2} = x_1(4k_0 + 2n_1 + n_0) \qquad (4.19)$$

formulát kapjuk, amit az eredeti $x(n)$ sorozattal kifejezve a következő összefüggések teljesülnek:

$$\begin{aligned}
x_1(k_0, 0, 0) &= x_0(0, 0, 0) + (-1)^{k_0} x_0(1, 0, 0) = x_1(0 + 4k_0) = \\
&= x(0) \pm x(4), \\
x_1(k_0, 0, 1) &= x_0(0, 0, 1) + (-1)^{k_0} x_0(1, 0, 1) = x_1(1 + 4k_0) = \\
&= x(1) \pm x(5), \\
x_1(k_0, 1, 0) &= x_0(0, 1, 0) + (-1)^{k_0} x_0(1, 1, 0) = x_1(2 + 4k_0) = \\
&= x(2) \pm x(6), \\
x_1(k_0, 1, 1) &= x_0(0, 1, 1) + (-1)^{k_0} x_0(1, 1, 1) = x_1(3 + 4k_0) = \\
&= x(3) \pm x(7),
\end{aligned} \qquad (4.20)$$

ahol a $\pm$ aktuális művelete k_0 értékétől függ.

Ezek után fejtsük ki a (4.17) sor belső n_1 szerinti összegzését, ekkor a DFT együtthatók

$$X(k) = \sum_{n_0=0}^{1} \left(\sum_{n_1=0}^{1} x_1(k_0, n_1, n_0) \epsilon_{-2(4k_2+2k_1+k_0)}^{8} \right) \epsilon_{-kn_0}^{8} \qquad (4.21)$$

alakú felírásából a fenti zárójelezett (n_1) összeghez tartozó x_2 sorozatot definiálhatunk a következő képpen:

$$x_2(k_0, k_1, n_0) = \sum_{n_1=0}^{1} x_1(k_0, n_1, n_0)(-i)^{(2k_1+k_0)n_1} = x_2(4k_0 + 2k_1 + n_0), \quad (4.22)$$

ahol kihasználtuk a $\epsilon_{-2}^{8} = -i$ és a $(-i)^{4k_2 n_1} = 1$ alakú azonosságokat, valamint, hogy most a $2n_1$ helyébe $2k_1$ lép a DFT-nek megfelelő összegzés miatt.

Az $N = 8$ esetnek megfelelően végül, az utolsó fokozat számításához definiáljuk az $x_3(n)$ sorozatot – az előzőeknek megfelelően – az $x_2(n)$ sorozat n_0 szerinti összegzésével:

$$x_3(k_0, k_1, k_2) = \sum_{n_0=0}^{1} x_2(k_0, k_1, n_0)\epsilon^8_{4k_2n_0}\epsilon^8_{(2k_1+k_0)n_0} = x_3(4k_0+2k_1+k_2), \quad (4.23)$$

ahonnan (4.21) figyelembevételével nyílvánvalóvá válik, hogy $x_3(n)$ sorozat nem más, mint a keresett $X(k)$ DFT együtthatók, mégpedig az $X(k) = X(k_2, k_1, k_0) = x_3(k_0, k_1, k_2)$ egyenlőség szerinti fordított bitképű indexekkel.

Összefoglalva tehát – az $N = 8$ esetben – a következő algoritmust vezettük le:

$$\begin{aligned} X(k_2, k_1, k_0) &= \sum_{n_0=0}^{1}\sum_{n_1=0}^{1}\sum_{n_2=0}^{1} x_0(n_2, n_1, n_0)\epsilon^8_{-4k_0n_2}\epsilon^8_{-2n_1(2k_1+k_0)}\epsilon^8_{-n_0(4k_2+2k_1+k_0)} = \\ &= \sum_{n_0=0}^{1}\sum_{n_1=0}^{1} x_1(k_0, n_1, n_0)\epsilon^8_{-2n_1(2k_1+k_0)}\epsilon^8_{-n_0(4k_2+2k_1+k_0)} = \\ &= \sum_{n_0=0}^{1} x_2(k_0, k_1, n_0)\epsilon^8_{-n_0(4k_2+2k_1+k_0)} = \\ &= x_3(k_0, k_1, k_2). \end{aligned} \quad (4.24)$$

Belátható, hogy a (4.22) és (4.23) szerint meghatározott $x_2(n)$ és $x_3(n)$ sorozatokra is a (4.20) összefüggésekhez hasonló pillangó-egyenleteket írhatunk fel. Például (4.22)-be $k_1 = 0$-át és $k_2 = 0$-át helyettesítve:

$$x_2(0,0,0) = x_1(0,0,0) + x_1(0,1,0) = x_2(0) = x_1(0) + x_1(2),$$

illetve

$$x_2(0,0,1) = x_1(0,0,1) + x_1(0,1,1) = x_2(1) = x_1(1) + x_1(3),$$

valamint hasonlóan (4.23)-ban $k_2 = k_1 = k_0 = 0$ indexekre:

$$x_3(0,0,0) = x_2(0,0,0) - x_2(0,0,1) = x_2(1) = x_2(0) - x_2(1).$$

Az elmodottak alapján azt gondolnánk, hogy a (4.24) alatti (R2 DIT2) algoritmus az – előző szakaszban bemutatott – R2 DIT1 FFT-től nagyban különbözik, azonban ez korántsem sincs így. Kicsit jobban összehasonlítva a két algoritmust észrevehető, hogy mindkettő a helyben számolásra épül, azonban ez utóbb bemutatott (4.24) eljárás esetében a bemenet lesz a normál és a kimenet lesz a fordított bitképű. Továbbá megfigyelhető, hogy mindkét FFT ugyanazokat az ϵ_{-k}^{N} komplex szorzó-konstansokat használja az egyes pillangó-műveletek elvégzése során. Megjegyezndő, hogy Cooley és Tukey az utóbbi R2 Dit2 algoritmust használta, vagyis ez tekinthető az összes többi FFT elődjének.

Az eddigi – R2 típusú – eljárásainkban feltettük, hogy $N = 2^L$ alakba írható, vagyis kettőhatvány. Az általánosításhoz tekintsük N következő felbontását:

$$N = \prod_{i=1}^{M} r_i, \quad r_i \in \mathbb{Z}, i \in \{1, 2, \ldots, M\}. \tag{4.25}$$

Ennek segítségével értelmezzük a (4.14) és (4.15) bináris felbontások általános formáját a következő kétváltozós

$$p(l, m) := \begin{cases} \prod_{i=m}^{l} r_i, & M \geq l \geq m, \\ 1, & l < m, \text{vagy } l > M \end{cases} \tag{4.26}$$

függvény bevezetésével. Ekkor ugyanis a $k_i = 0, 1, \ldots, r_i - 1$ és az $n_i = 0, 1, \ldots, r_{M-i+1} - 1$ egész számok felhasználásával a

$$k = \sum_{i=1}^{M} k_i p(i-1, 1), \tag{4.27}$$

illetve

$$n = \sum_{i=1}^{M} n_i p(M, M-i+2), \tag{4.28}$$

egyértelmű felbontásokkal adhatók meg. Könnyen belátható, hogy ezek a definíciók valóban tekinthetők a (4.15), illetve a (4.14) bináris felírások általánosításának. Tekintsük például az $N = 8$ esethez tartozó felbontást. Ekkor nyilván $N = 2 \cdot 2 \cdot 2$, azaz $M = 3$ és $r_i = 2$ ($\forall i \in \{1, 2, 3\}$), melynek következtében $p(0, 1) = p(3, 4) = 1$, $p(1, 1) = p(3, 3) = 2$ és $p(2, 1) = p(3, 2) = 4$ értékek éppen a kívánt bináris felbontást adják.

A továbbiakhoz tekintsük az $x(n)$ sorozatot egy M-dimenziós $x(n) = x_0(n_1, n_2, \ldots, n_M)$ sorozatnak. Ekkor felírható a DFT együtthatók meghatározására egy – a (4.17) alatti bináris felbontáshoz hasonló – többszörös összegzési formula, nevezetesen $X(k)$ az

$$X(k) = \sum_{n_1=0}^{r_M-1} \cdots \sum_{n_M=0}^{r_1-1} x_0(n_1, n_2, \ldots, n_M) \epsilon_{-nk}^{N} \tag{4.29}$$

egyenlettel számolható. A korábbiakhoz hasonlóan, most is helyettesítsük be az n és k – azaz nk – helyére a (4.28) és (4.27) alatti felbontásokat néhány apró, de igen fontos észrevétellel. Nevezetesen, ismeretes, hogy ϵ^N az N-re periodikus, ezért (nk) helyett számolhatunk $(nk)\ mod\ N$ értékkel. Ugyanakkor a $p(1,M) = N$ miatt a $p(M, M-i+2)p(l-1,1)$ szorzat értéke az N egész számú többszöröse, ha $l \geq M-i+2$. Ezek után az (nk) szorzat a

$$\begin{aligned} nk &= \left(\sum_{i=1}^{M} n_i p(M, M-i+2)\right)\left(\sum_{l=1}^{M} k_l p(l-1,1)\right) = \\ &= \sum_{i=1}^{M}\left(n_i p(M, M-i+2)\sum_{l=1}^{M-i+1} k_l p(l-1,1)\right) \end{aligned} \tag{4.30}$$

egyszerűsített alakra hozható, ahol észrevehetően i szerinti kifejtést használtunk.

A most meghatározott – (4.30) alatti – (nk) kifejtést a (4.29) egyenletbe behelyettesítve definiálhatjuk a legbelső – n_M index szerinti – összegzést egy $x_1(n) = x_1(n_1, n_2, \ldots, n_{M-1}, k_1)$ sorozatként a

$$\begin{aligned} x_1(n_1, n_2, \ldots, n_{M-1}, k_1) &= \sum_{n_M=0}^{r_1-1} x_0(n_1, \ldots, n_M)\epsilon^N_{-n_M k_1 p(M,2)} = \\ &= \sum_{n_M=0}^{r_1-1} x_0(n_1, \ldots, n_M)\epsilon^{r_1}_{-n_M k_1} \end{aligned} \tag{4.31}$$

egyenlet formájában, ahol kihasználtuk, hogy a k_1 és n_M indexek mindig együtt szerepelnek, tehát felcserélhetőek az $x_1(n)$ sorozatban, valamint a $p(M,2) = N/r_1$ következményeként előálló $\epsilon^N_{p(M,2)} = \epsilon^{r_1}$ egyszerűsítést is alkalmaztuk.

Hasonlóan az R2 DIT2 esethez most is a következő lépés egy $x_2(n) = x_2(n_1, n_2, \ldots, n_{M-2}, k_2, k_1)$ sorozat definiálása, kifejtve a (4.29) sor soron következő – n_{M-1} index szerinti – legbelső összegzését $x_1(n)$ segítségével:

$$x_2(n_1, n_2, \ldots, n_{M-2}, k_2, k_1) = \sum_{n_{M-1}}^{r_2-1}\left(x_1(n_1, n_2, \ldots, n_{M-1}, k_1)\epsilon^{r_1 r_2}_{-k_1 n_{M-1}}\right)\epsilon^{r_2}_{-k_2 n_{M-1}}. \tag{4.32}$$

Mivel a N számot M darab r_i szorzó-tényezőre bontottuk, ezért ezt az eljárást M lépésen keresztül rekurzívan folytatni tudjuk, melynek végén éppen a kívánt $X(k) = x_M(k_M, k_{M-1}, \ldots, k_2, k_1)$ DFT együtthatókat kapjuk.

Az algoritmus egy közbülső m-edik lépését szemlélteti a következő formula, a $k_M^* = \sum_{i=1}^{m-1} k_i p(i-1,1)$ egyszerűsített jelölés bevezetésével:

$$x_m(n_1,\ldots,n_{M-m},k_m,\ldots,k_1) = \sum_{n_{M-m+1}=0}^{r_m-1} \Big(x_{m-1}(n_1,\ldots,n_{M-m+1},$$

$$k_{m-1},\ldots,k_1)\epsilon_{-n_{M-m+1}k_m^*}^{p(m,1)}\Big)\epsilon_{-k_m n_{M-m+1}}^{r_m}, \tag{4.33}$$

ahol a közbenső zárójelben lévő sorozat r_m pontos pillangóval megvalósított DFT-je, éppen az x_m sorozatot adja. Az x_m sorozatot pedig az x_{m-1} sorozatból tulajdonképpen egy *elforgatással* – exponenciális $\epsilon_{-n_{M-m+1}k_m^*}^{p(m:1)}$ szorzattal – határozzuk meg. Ezt a mennyiséget a szakirodalomban *előforgatási tényezőnek* [1] is szokták nevezni.

A fent definiált módszer alkalmat ad más FFT eljárások – különböző algebrai módon történő – származtatására, melynek során a levezetésnél használt M-dimenziós tömböket egydimenziós sorozatokká kell transzformálni.

4.3. Algoritmusok osztályozásai

Amióta 1965-ben Cooley és Tukey megalkották az – előző 4.2. fejezetben részletesebben tárgyalt – első FFT algoritmust, nagyon sok, lényegében hasonló algoritmus született a DFT kiszámításának gyorsításásra. Ezeket az FFT-ket különböző tulajdonságaik alapján csoportba soroljuk.

Az egyik legfontosabb osztályozási szempont az, hogy az időtartományi skálán kezdjük-e el a lebontási folyamatot, vagy a frekvencia-tartomány oldalán. Amennyiben először a bemenet mintáit – tehát az időtartománybeli értékeket – fogjuk össze kisebb csoportokba a kisebb pontszámú DFT számításához, úgy – a már két algorimus által is bemutatott – *DIT* (Decimation In Time) algoritmusok osztályáról beszélünk. Ha viszont a frekvencia-tartomány felől indítjuk el a lebontó folyamatot, akkor egy másik, úgynevezett *DIF* (Decimation In Frequency) [1] algoritmusok osztályához jutunk.

Másik osztályozási szempont a bemeneti, illetve kimeneti sorozatok bitjeinek sorrendisége. Mint láthattuk, ez a sorrend lehet *természetes* (vagy szekvenciális), illetve *fordított bitképű* (vagy nemszekvenciális). A sorozatok bitjeinek a megfordításához a (4.13) alatt bevezetett *bitfordító-transzformáció* nagyon jól használható.

Végül a harmadik osztályozási szempont a *pillangók, különböző fokozataihoz tartozó, geometriai elhelyezkedésével* hozható kapcsolatba. Már beszéltünk az úgynevezett helyben számolásról, amikor az egyes – különböző fázisban lévő – pillangó-műveletek ugyanazt a memória-rekeszt használják.

A hatásgráfot úgy kell elképzelni, hogy a gráf jobb, illetve bal oldalán egymás alatt fel vannak sorakoztatva a bemeneti, illetve a kimenetei sorozatok elemei, e két sorozat között pedig – a különböző fázisokat szimbolizáló – csomópontok vannak, a fázisok számának megfelelő számú oszloba rendezve. Az egyes fázisok csomópontjai aszerint vannak összekötve, hogy éppen melyik két kiindulási együttható segítségével végzett pillangó-művelet melyik két együtthatót eredményezi. Helyben számolás esetén – a hatásgráfon – az egyes pillangó-műveletek bemeneti és kimeneti pontjai egy képzeletbeli vízszintes egyenesenen helyezkednek el. Ezen túlmenően vannak olyan FFT algoritmusok, melyeknél a pillangók bemeneti, illetve kimeneti pontjai minden fokozatban azonos vízszintes magasságú pontokhoz csatlakoznak. Az ilyen osztályokba tartozó algoritmusok az *azonos bemeneti*, illetve *azonos kimeneti geometriájú* algorimusok. Ezek közül a legfontosabb FFT típus az úgynevezett *izogeometrikus* FFT, amelynél minden pillangó gometriája teljesen azonos, azaz minden pillangó – az egymás utáni fokozatokban – azonos magasságú vízszintes vonalról indul és ettől eltérő, de ugyancsak megegyező magasságú vízszintes vonalra érkezik [1].

A most bemutatott szempontok alapján 8 DIT és 8 DIF típusú alap algoritmust tudunk megkülönböztetni, amelyek apróbb változtatásával szinte az összes FFT algoritmust magukban foglalják. A továbbiakban a 4.2 szakaszban tárgyalt algebrai általánosítás segítségével bevezetjük az egyes alap algoritmusok formális definícióját.

Először a DIT algoritmusok általánosabb változatait mutatjuk be. Ehhez jelöljük l_m-mel ($m \in \{0, \ldots, M\}$) azt az egydimenziós indexet, amelyet az adott algoritmus az m-medik lépésben az n_i és k_i ($i \in \{1, \ldots, M\}$) indexekből állít elő. A definícióból következik, hogy az $m = 0$ esteben visszakapjuk a bemeneti, míg $m = M$ esetén a kimeneti DFT sorozathoz tartozó indexeket. Mivel az egyes DIT algoritmusok, tulajdonképpen ezen l_m indexek előállításának módjában különböznek egymástól, ezért most csak ezeket mutatjuk be.

Mint láthattuk, a DIT1 és DIT2 algoritmusok a helyben számolás megkötésére épülnek. Ekkor a $p(s, t)$ függvény előállításakor az s és t értékek függetlenek m fokozatszámtól. A (4.28) és (4.27) előállítások figyelembe vételével a Cochran által bevezetett algoritmus a következő

$$l_m = \sum_{i=1}^{m} k_i p(i-1, 1) + \sum_{i=m+1}^{M} n_{M-i+1} p(i-1, 1) \tag{4.34}$$

indexelésre épül, ahonnan $m = 0$ és $m = M$ helyettesítéssel visszakapjuk az említett speciális eseteket:

$$
\begin{aligned}
l_0 &= \sum_{i=1}^{M} n_{M-i+1} p(i-1, 1) \simeq (n_M, n_{M-1}, \ldots, n_1) && (4.35) \\
l_M &= \sum_{i=1}^{M} k_i p(i-1, 1) \simeq (k_M, k_{M-1}, \ldots, k_1). && (4.36)
\end{aligned}
$$

Ez utóbbi esetekből világosan látszik, hogy a bemenetre a fordított bitképű sorozatnak kell kerűlni, a kimeneten pedig a természetes sorrendben jelennek meg a DFT együtthatók.

A DIT2 – Cooley-Tukey féle – algoritmushoz tartozó egydimenziós index alőállítása az

$$
l_m = \sum_{i=1}^{M-m} n_i p(M, M-i+2) + \sum_{i=M-m+1}^{M} k_{M-i+1} p(M, M-i+2) \qquad (4.37)
$$

alakot ölti, ahol most a bemenet ($l = 0$) bitképe lesz természetes sorrendű és a kimeneté ($m = M$) lesz fordított.

A további DIT algoritmus már nem használja a helyben számolás elvét, éppen ezért $p(s, t)$ meghatározásában az s és t értékek már nem függetlenek m-től. A DIT3 indexezését a következő alakban állíthatjuk elő:

$$
l_m = \sum_{i=1}^{m} k_i p(m, i+1) + \sum_{i=1}^{M-m} n_i p(M-i, 1), \qquad (4.38)
$$

ahonnan – behelyettesítés után – nyilvánvaló, hogy mind a bemenet, mind a kimenet bitképe fordított sorrendű lesz.

A DIT4 algoritmus is Cochran nevéhez fűződik, az indexezés

$$
l_m = \sum_{i=1}^{M-m} n_i p(M, M-i+2) + p(M : m+1) \sum_{i=1}^{m} k_i p(i-1, 1) \qquad (4.39)
$$

alakú, ahol most a bemeneti, illetve a kimeneti sorozatok indexezése egyaránt természetes sorrendű. Amennyiben az előállító $r_1, \ldots, r_M$ radix számok megegyeznek, úgy az egymás után következő fokozatokban a pillangók kimeneti geometriája megegyezik.

A DIT5 FFT egydimenziós indexe az

$$
l_m = \sum_{i=1}^{M-m} n_i p(M-i, m+1) + \sum_{i=1}^{m} k_i p(M, i+1) \qquad (4.40)
$$

alakban adható meg, melynek megfelelően mind a bemenet, mind a kimenet sorrendje – előnytelen – digitképű.

A DIT6 algoritmust Gentleman és Sande publikálták először. Ekkor az indexezés az

$$l_m = \sum_{i=1}^{m} k_i p(i-1,1) + p(m,1) \sum_{i=1}^{M-m} n_i p(M, M-i+2) \tag{4.41}$$

alakba írható, ahol a kimenet is és a bemenet is természetes sorrendű. Ha az összes radix szám megegyezik, akkor minden fokozatban azonos bemeneti geometriával rendelkező pillangókkal számol az algoritmus.

Az úgynevezett Singleton-féle DIT7 esetében az egydimenziós indexelés

$$l_m = \sum_{i=1}^{M-m} n_i p(M-i, m+1) + p(M, m+1) \sum_{i=1}^{m} k_i p(i-1,1) \tag{4.42}$$

alakú. Most a bemenet sorrendje fordított, míg a kimeneté természetes bitképű. Az azonos radix számokhoz tartozó előállítás esetében a már említett izogeometrikus tulajdonság mintapéldáját kapjuk, amely a soros számítógépes feldolgozás szempontjából komoly előnyökkel jár.

A DIT8 algoritmus hatásgráfja is izogeometrikus, ahol az indexezés az

$$l_m = \sum_{i=1}^{m} k_i p(m, i+1) + p(m,1) \sum_{i=1}^{M-m} n_i p(M, M-i+1) \tag{4.43}$$

alakot ölti. A bemenet természetes sorrendű, a kimenet pedig fordított bitképű.

Most a DIF, azaz a frekvencia decimálásán alapuló algoritmusok származtatását mutatjuk be. Ebben az esetben az $X(k)$ DFT együtthatók mintáit kell csoportosítani. A páros N pontszámú DIF algoritmusok első lépése – a $k = 0, 1, \ldots, N-1$ értékek mellett – az

$$\begin{aligned} X(k) &= \sum_{n=0}^{N/2-1} x(n)\epsilon_{-nk}^{N} + \sum_{n=N/2+1}^{N} x(n)\epsilon_{-nk}^{N} = \\ &= \sum_{n=0}^{N/2-1} \left(x(n) + \epsilon_{-kN/2}^{N} x\left(n + \frac{N}{2}\right)\right) \epsilon_{-nk}^{N} \end{aligned} \tag{4.44}$$

összefüggésre épül. Az $X(k)$ együtthatókat – figyelembe véve az $\epsilon^N_{-kN/2} = (-1)^k$ azonosságot – szétbonthatjuk a páros, illetve páratlan indexű együtthatók összegére:

$$X(2k) = \sum_{n=0}^{N/2-1} \left(x(n) + x\left(n + \frac{N}{2}\right)\right) \epsilon^N_{-2kn}, \quad \text{illetve} \tag{4.45}$$

$$X(2k+1) = \sum_{n=0}^{N/2-1} \left(\left(x(n) - x\left(n + \frac{N}{2}\right)\right) \epsilon^N_{-n}\right) \epsilon^N_{-2kn}, \tag{4.46}$$

ahol $k = 0, 1, \ldots, N/2 - 1$. Ennek következménye képpen az N-pontos DFT kiszámítását felbontottuk két $N/2$-pontos DFT elvégzésére. Ez a felbontás tovább folytatható, egészen a kétpontos DFT műveletek elvégzéséig. A kétpontos alapműveletekhez tartozó pillangó műveletek általános formája a – DIT FFT-k (4.5) és (4.6) alatt bevezetett pillangó-műveletekhez hasonló – következő formát ölti:

$$X_1 = x_1 + x_2, \tag{4.47}$$

$$X_2 = (x_1 - x_2)\epsilon^N_{-k}. \tag{4.48}$$

Az összefüggés a két különböző pillangó hatásgráfja között az, hogy az egyikből a másik az irányok, illetve a bemenet és kimenet szerepének megcserélésével előállítható. Azt mondhatjuk, hogy az egyik pillangó-elrendezés a másik elrendezés *transzponáltja*.

Könnyen ellenőrizhető az is, hogy az a transzponált kapcsolat a DIT és DIF FFT-k egészére vonatkozóan is fennáll [1]. Ennek következtében, ahelyett, hogy a DIF algoritmusok családját a DIT1,...,DIT8 algoritmusokhoz hasonló módon levezetnénk, megadjuk az egyes DIF FFT-k transzponáltpárjait. Tehát párosítjuk az egyes DIF algoritmusokat, azzal a DIT algoritmussal, melynek mátrix értelemben vett transzponáltja megegyezik az adott DIF eredő hatásgráfjával.

Az algebrai megadási módhoz bontsuk fel – a (4.30)-hoz hasonló módon – az nk szorzat helyett a kn szorzatot:

$$kn = \sum_{m=1}^{M} \left(k_m p(m-1, 1) \sum_{i=1}^{M-m+1} n_I p(M, M-i+2)\right). \tag{4.49}$$

Ennek segítségével – (4.31) helyett – az $x_1(n)$ sorozatra a

$$x_1(n_1, \ldots, n_{M-1}, k_1) = \sum_{n_M=0}^{r_1-1} x_0(n_1, \ldots, n_M)\epsilon^{r_1}_{-k_1 n_M}\epsilon^N_{-k_1 n^*_M} \tag{4.50}$$

írható, ahol $n_M^* = \sum_{i=1}^{M-1} n_i p(M, M-i+2)$. Ezt az eljárást rekurzív módon tovább folytatva, most az m-medik lépésre – az $n_M^* = \sum_{i=1}^{M-m} n_i p(M, M-i+2)$ egyszerűsített jelölés bevezetésével – az

$$x_m(n_1, \ldots, n_{M-m}, k_m, \ldots, k_1) = \sum_{n_{M-m+1}=0}^{r_m-1} \big(x_{m-1}(n_1, \ldots, n_{M-m+1},$$

$$k_{m-1}, \ldots, k_1)\epsilon_{-k_m n_{M-m+1}}^{r_m}\big)\epsilon_{-k_m n_m^*}^{p(M,m)} \tag{4.51}$$

formula írható fel.

A most látott k szerinti felbontás sgítségével is definiálhatjuk az egyes DIF algoritmusokhoz tartozó l_m index előállítást. Az alábbi felsorolás az FFT algoritmusok transzponált kapcsolatatit hivatott bemutatni: DIT1-DIF2, DIT2-DIF1, DIT3-DIF5, DIT4-DIF6, DIT5-DIF3, DIT6-DIF4, DIT7-DIF8 és DIT8-DIF7. Természetesen az egyes DIF algoritmusok tulajdonágai rendre megegyeznek a hozzájuk tartozó transzponált-párok tulajdonságaival, ezért azok ismertetését most kihagyjuk.

A DFT és inverz DFT (3.21), illetve (3.23) alatti definíciójának következménye, hogy az FFT algoritmusok az inverz FFT (IFFT) operáció közvetlen elvégzésére is alkalmasak. Ugyanis IFFT esetében a komplex ϵ_{-1}^N együttható helyett annak komplex konjugáltjával, tehát $\epsilon_1^N = \epsilon^N$-nel kell szorozni, majd a kapott eredményt N-nel osztani kell. Azonban megadható egy olyan formula is az IFFT meghatározására, amelynél magán az FFT algoritmuson nem kell változtatni:

$$\overline{x}(n) = \overline{\left(\frac{1}{N}\sum_{k=0}^{N-1} X(k)\epsilon_{kn}^N\right)} = \frac{1}{N}\sum_{k=0}^{N-1} \overline{X}(k)\epsilon_{kn}^N\ . \tag{4.52}$$

Azonban, mint látható, ekkor a frekvencia oldalon a $x(n)$ sorozat komplex konjugáltját kapjuk.

4.4. Gyors konvolúció és az FFT

A 3. fejezetben a z-transzformáció bevezetésekor már bemutattuk, hogy egy líneáris, időinvariáns, diszkrét idejű rendszer miként épül fel. Nevezetesen, (3.25) szerint, bármely ilyen típusú jel felírható a $\delta(n)$ egység impulzus függvénnyel való konvolúciós alakban, továbbá a rendszer kauzalitása miatt az előállító sor véges összegzésbe megy át. Egy véges hosszúságú N-pontos DFT – azaz FFT – segítségével analizálni tudjuk az időtartományú jelsorozatot és egyszerűen meg tudjuk határozni, hogy milyen periodicitású szinuszos és koszinuszos sorozatokból lehet összerakni, illetve visszaállítani az adott jelet. Továbbá az inverz DFT (illetve FFT) egyszerűsége miatt a frekvencia-tartományból az időtartományba történő transzformáció is könnyen elvégezhető. A Fourier-transzformáció és konvolúció (2.19) alatt bevezetett kapcsolata miatt a konvolúció meghatározása a direkt módon történő előállíásnál gyorsabban is elvégezhető a DFT segítségével. És mivel a DFT – illetve inverz DFT – meghatározására léteznek gyors FFT algoritmusok, ez a másik fajta, úgynevezett *gyors konvolúciós* előállítás méltán válhatott a jelfeldolgozás egyik alappillérévé.

Legyen $h(n)$ az N_1 hosszúságú véges impulzusválasz, $x(n)$ pedig egy N_2 hosszúságú diszkrét jelsorozat. Ekkor az $y(n)$ kimenő jelsorozat az

$$y(n) = x(n) * h(n) = \sum_{k=0}^{n} x(k)h(n-k) = \sum_{k=0}^{n} h(k)x(n-k) \tag{4.53}$$

konvolúcióval számítva maximum $N_0 = N_1 + N_2 - 1$ hosszúságú lehet, azaz legfeljebb ennyi zérustól különböző mintát tartalmazhat. Ekkor $y(n)$ kimenő jelsorozat a gyors konvolúcióval a következő algoritmussal állítható elő:

1. *lépés*:
   A $h(n)$ sorozatot $N_2 - 1$, az x(n) sorozatot pedig $N_1 - 1$ zérussal kiegészítjük, hogy mindkét jel N_0 hosszúságú legyen.

2. *lépés*:
   Meghatározzuk a $h(n)$ véges impulzusválaszhoz tartozó $H(k)$ N_0-pontos DFT együtthatókat a (3.21) szerinti

   $$H(k) = \sum_{n=0}^{N_0-1} h(n)\epsilon_{-nk}^{N_0}, \quad k = 0, 1, \ldots, N_0 - 1 \tag{4.54}$$

   összefüggéssel, ahol $\epsilon_{-nk}^{N_0} = e^{-i2\pi kn/N_0}$. Természetesen a DFT meghatározása FFT segítségével történik.

3. *lépés*:
   Meghatározzuk az $x(n)$ diszkrét jelhez tartozó $X(k)$ N_0-pontos DFT együtthatókat a szintén (3.21) szerinti

   $$Z(k) = \sum_{n=0}^{N_0-1} x(n)\epsilon_{-nk}^{N_0}, \quad k = 0, 1, \ldots, N_0 - 1 \tag{4.55}$$

   formulával, ahol a kiszámítás most is az FFT segítségével történik.

4. *lépés*:
   Képezzük a (2.19) szerinti $Y(k) = H(k)X(k)$ szorzatokat $k = 0, 1, \dots, N-1$ esetekre, tehát a két sorozat konvolúcióját a DFT sorozataik szorzatával hajtjuk végre, ahol $Y(k)$ éppen az $y(n)$ DFT-je (illetve FFT-je).

5. *lépés*:
   Egy N_0 pontos inverz DFT-t – azaz IFFT-t – számolunk $Y(k)$ együtthatókból a (3.23) szerinti

$$y(n) = \frac{1}{N_0} \sum_{k=0}^{N_0-1} Y(k)\epsilon_{nk}^{N_0}, \quad n = 1, 2, \dots, N_0 - 1 \tag{4.56}$$

   formula segítségével.

Általában a bemenő és a kimenő jelek komplex alakúak, tehát a műveleteket is komplex szorzásként, illetve összeadásként értelmezzük. Ezek a műveletek természetesen a (4.2) összefüggések alapján átalakíthatók valós operációkká. Mint ahogyan azt a 4.1. szakasz bevezetésében tárgyaltuk a DFT direkt módon történő kiszámításának műveletigénye $\mathcal{O}(N_0^2)$. Az FFT alkalmazásával ez lecsökkenthető $\mathcal{O}(N_0 \log_2 N_0)$-ra. Most vizsgáljuk meg a gyors konvolúció egyes lépéseinek műveletigényét. Az első lépés műveletigénye $N_{min} := \min(N_1, N_2)$, mivel legfeljebb ennyi zérussal kell feltölteni. A második és harmadik lépések – az FFT alkalmazásával – $\mathcal{O}(N_0 \log_2 N_0)$ idő alatt elvégezhetők. A negyedik lépés egyetlen komplex szorzásból áll, ezért annak műveletigénye konstans 1. Végül az ötödik lépés inverz Fourier-transzformációja szintén elvégezhető $\mathcal{O}(N_0 \log_2 N_0)$ idő alatt, hiszen az IFFT az FFT-től csak egy előjelben és egy komplex konjugáltban jelentkezik. Összefoglalva tehát az algoritmus össz-műveletigénye:

$$\mathcal{O}(N_{min}) + \mathcal{O}(N_0 \log_2 N_0) + \mathcal{O}(N_0 \log_2 N_0) + 1 + \mathcal{O}(N_0 \log_2 N_0) = \mathcal{O}(N_0 \log_2 N_0),$$

ami – mint már láttuk – azt jelenti, hogy $N_0 = 1024$ esetén a gyors konvolúció alkalmazása 100-szoros javulást eredményez.

5. fejezet

Waveletek

5.1. Waveletek szerkesztése

A *waveletek* olyan függvényosztályoknak a gyüjtőnevei, amelyek elget tesznek bizonyos feltételeknek. A wavelet szó az angol hullám (wave) szóból ered, sugallva azt a kívánságot, hogy a szóban forgó függvények hullámozzanak a x-tengely felett és alatt, azaz integráljuk legyen nulla. Egy másik követelmény, hogy legyenek jól lokalizálhatók, úgy a frekvencia-tartományban, mint az időskálán. Természetesen vannak még további elvárások a waveletekkel szemben, melyek elsősorban technikai jellegűek azért, hogy a velük megvalósított transzformációk gyorsan elvégezhetők legyenek.

A waveletek ugyanakkor tekinthetőek a folytonos időben bázis függvényeknek is. A bázis líneárisan független függvények halmaza, amelyek előállítják az adott téren értelmezett $f(x)$ függvényeket, azaz

$$f(x) = \sum_{n,k} b_{nk} \psi_k^n(x). \tag{5.1}$$

A speciális tulajdonsága azonban a $\psi_k^n(x)$ waveleteknek az, hogy előállíthatók egy úgynevezett $\psi(x)$ *anya-waveletből*. Ez az alapfüggvény egy kisebb hullám (egy pulzus). Általában az $x = 0$ időben indulnak és az $x = N$ időben véget is érnek. Ennek megfelelően az eltolt $\psi_0^k(x) = \psi(x - k)$ waveletek az $x = k$ időpontban indulnak és az $x = k + N$ időpontban érnek véget, más szóhasználattal élve, a ψ_0^k függvények *tartója* a $[k, N + k]$ intervallum. Az átskálázott $\psi_n^0(x) = \psi(2^n x)$ waveletek indulási időpontja az $x = 0$, de már az $x = N/2^n$ időpontban véget érnek, mintegy összenyomva a gráfjukat egy 2^n faktorral (tehát a tartójuk a $[0, N/2^n]$ intervallum).

A wavleteknek számos fajtája létezik, ebben a szakaszban az úgynevezett *affin waveletekkel* foglalkozunk [8]. Azaz kiindulunk egy $\psi \in L^2(\mathbb{R})$ alapfüggvényből és azt vizsgáljuk, hogy milyen feltétel mellett lesz az alapfüggvényből

– az 2. fejezetben bevezetett – transzlációval és dilatációval származtatott

$$\psi_k^n(x) := 2^{n/2}\psi(2^n x - k), \quad x \in \mathbb{R}, k, n \in \mathbb{Z} \tag{5.2}$$

függvényrendszer ortonormált a szokásos $L^2(\mathbb{R})$-beli

$$< f, g >= \int_{\mathbb{R}} f(x)\overline{g(x)}\, dx, \quad f, g \in L^2(\mathbb{R}) \tag{5.3}$$

skaláris szorzatra nézve. A továbbiakban feltesszük, hogy az anya-waveleteknek az $L^2(\mathbb{R})$ téren vett normájuk 1.

A legegyszerűbb ilyen wavelet-konstrukció a

$$\psi(x) := h(x) = \begin{cases} 1, & 0 \leq x < 1/2, \\ -1, & 1/2 \leq x < 1, \\ 0, & x \notin \mathbb{R}\ [0, 1) \end{cases} \tag{5.4}$$

anya-wavelettől képzett

$$h_k^n(x) = 2^{n/2}h(2^n x - k), \quad x \in \mathbb{R}, k, n \in \mathbb{Z} \tag{5.5}$$

Haar-féle ortonormált rendszer. Könnyen igazolható, hogy a fenti – sok vonatkozásban is kitüntetett – rendszer valóban ortonormált. Hasonló waveletek szerkesztéséhez felhasználjuk a Fourier-analízis – korábban már ismertetett – eszköztárát.

A 6.3. szakaszban a waveletek és szűrőkészletek közötti kapcsolat bemutatásánál látni fogjuk, hogy az úgynevezett *felüláteresztő szűrők (highpass filter)* azok, amelyek a $\psi(x)$ wavelethez vezetnek, míg az *alulátersztő szűrők (lowpass filter)* a – nemsokára bemutatásra kerülő – $\varphi(x)$ skála függvényeket állítják elő. A legtöbb konstrukcióban az aluláteresztő szűrő jön előbb, tehát a skála függvényt előbb határozzuk meg, mint magát a waveletet. Valójában a skála függvény (folytonos időben) nem más, mint az aluláteresztő szűrők végtelen sokszori egymás utáni alkalmazása, ahol a szűrőket minden egyes iterációs lépésben újraskálázunk. A waveletet viszont a $\varphi(x)$ függvényből a felülátersztő szűrő egyszeri alkalmazásával kapunk meg.

5.2. Multirezolúció

Adott felbontásán egy jelnek vagy képnek, a bázisok – a jelek egy halmazára nézve – nem mások, mint a $\varphi(2^n x - k)$ alakú skála függvények. Az n-nedik szinten az időbeli lépések mértéke 2^{-n}. Az új részleteket (details [2]) ugyanezen szinten a $\psi(2^b x - k)$ alakú waveletek szolgáltatják. A sima jel plussz a részletek, tehát a φ és ψ függvények együttese, az $n+1$-edik – finomabb – szint *multirezolúcióját* alkotják. A sima függvények nem mások, mint lokális átlagok, melyek a skála függvényekből származnak, míg a waveletekből származó részletek a lokális különbségeket jelentik, azaz:

$$\text{jel az } n. \text{ szinten } + \text{ részlet az } n. \text{ szinten } = \text{ jel az } n+1. \text{ szinten.}$$

Szemléletesen ez jelenti a multirezolúcióját egy jelnek. Ha mindezt alkalmazzuk az összes jelünkre, kapjuk az adott terek egy multirezolúcióját, azaz:

$$V_n \text{ (skála függvények tere)} \oplus W_n \text{ (wavelet függvények tere)} = V_{n+1},$$

ahol az indexek a szint-számot jelölik. A multirezolúció ötlete képezi tulajdonképpen a wavelet-analízis alapját. A következőkben megadjuk a most bevezetett felbontás pontos matematikai definícióját.

A cél tehát a függvénytér alterekre való felbontása. Ehhez először bevezetjük az ortonormált bázis fogalmának egy általánosítását, az úgynevezett *Riesz-bázisokat* [8]. Jelöljük ℓ_0^2-tel azoknak az ℓ^2-beli konvergens sorozatoknak a halmazát, amelynek csak véges sok tagja különbözik nullától. Egy $\varphi_k \in L^2(\mathbb{R})$ $(k \in \mathbb{Z})$ függvényrendszer Riesz-bázis, ha léteznek olyan $0 < m \leq M < \infty$ számok, melyekre minden $c = (c_k, k \in \mathbb{Z}) \in \ell_0^2$ sorozatra a

$$m||c||_{\ell^2} \leq ||\sum_{k\in\mathbb{Z}} c_k\varphi_k||_{L^2(\mathbb{R})} \leq M||c||_{\ell^2} \tag{5.6}$$

egyenlőtlenség teljesül. Könnyen ellenőrizhető, hogy a fenti definíció valóban tekinthető az ortonormált bázis általánosításának, ugyanis, ha a φ_k rendszer ortonormált, akkor bármely $c \in \ell_0^2$ sorozatra fennáll a

$$||\sum_{k\in\mathbb{Z}} c_k\varphi_k||_{L^2(\mathbb{R})} = ||c||_{\ell^2}$$

egyenlőség. Továbbá, igazolható, hogy a definícióban szereplő végtelen sor a tér normájában konvergens és az ilyen alakú összegek az $L^2(\mathbb{R})$ tér egy zárt alterét alkotják. A továbbiakban olyan altereket vizsgálunk, amelyeket

$$\varphi_k := \tau_k\varphi, \quad \varphi \in L^2(\mathbb{R}), ||\varphi||_{L^2(\mathbb{R})} = 1, k \in \mathbb{Z} \tag{5.7}$$

alakú, egész transzlációkkal szemben invariáns Riesz-bázisok generálnak. Bizonyítható, hogy ha φ_k rendszer Riesz-bázis (ortonormált), akkor minden n egész szám esetén a

$$\varphi_k^n(x) := 2^{n/2}\varphi(2^n x - k), \quad x \in \mathbb{R}$$

alakú függvényrendszer is az.

Mindezek segítségével már definiálható az úgynevezett *multirezolúciós (MR) felbontás*, amelyet azon $V_n \subseteq L^2(\mathbb{R})$ zárt alterek sorozata alkot, amelyekre a következő feltételek teljesülnek:

1. $V_n \subseteq V_{n+1}$, $n \in \mathbb{Z}$, (monotonitási feltétel)
2. $\cup_{n\in\mathbb{Z}} V_n$ mindenütt sűrű az $L^2(\mathbb{R})$ tér pontjaiban, (sűrűségi feltétel)
3. $\cap_{n\in\mathbb{Z}} V_n = \{0\}$, (szeparációs feltétel)
4. $\forall n \in \mathbb{Z}$ esetén $f \in V_n$, akkor és csak akkor, ha $\delta_2 f \in V_{n+1}$ (skálázási feltétel) és
5. a V_0 teret egy (5.7) alakú Riesz-bázis generálja.

Jelöljük V_n^φ-vel a (5.7) Riesz-bázis által generált alteret:

$$V_n^\varphi := \left\{ f := \sum_{k\in\mathbb{Z}} c_k \varphi_k^n \,\middle|\, c = (c_k, k \in \mathbb{Z}) \in \ell^2 \right\}, \tag{5.8}$$

amely terek bázisfüggvényeit *skála függvényeknek* nevezzük. Bizonyítható, hogy a szeparációs feltétel a többi következménye, ezért azzal a továbbiakban nem foglalkozunk. A két utolsó feltétel alapján igazolható, hogy a φ_k^n $(k \in \mathbb{Z})$ rendszer a V_n-nek egy Riesz-bázisa, következésképpen a V_n altér a $\tau_{k2^{-n}}$ alakú transzlációkkal szemben invariáns. A skálázási feltétel interpretációja, hogy a δ_2 dilatációval bármely V_n altérről áttérhetünk az eggyel magasabb indexű V_{n+1} altérre, azaz a *következő szintre*. Mindezek figyelembevételével a φ függvény tehát egyértelműen meghatározza a fenti MR-felbontást.

Példaként bemutatjuk a legegyszerűbb MR-felbontást, amely a $[0, 1)$ intervallum $\varphi := \chi_{[0,1)}$ karakterisztikus függvényéből kiindulva szerkeszthető. Ekkor V_n^φ olyan $L^2(\mathbb{R})$-beli függvényekből áll, amelyek a $[k2^{-n}, (k+1)2^{-n})$ $(k \in \mathbb{Z})$ diadikus intervallumokon állandók. Nyilvánvaló, hogy $V_n^\varphi \subseteq V_{n+1}^\varphi$ és φ_k^n a V_n^φ-nek egy ortonormált bázisa.

Problémát okoz azonban az a tény, hogy két különböző szinten lévő bázis elemek nem ortogonálisak egymásra. Jelöljük W_n^φ-vel a V_n^φ altérnek V_{n+1}^φ-re vonatkozó ortogonális kiegészítőjét (komplementerét), azaz $V_{n+1}^\varphi = V_n^\varphi \oplus W_n^\varphi$. Ekkor a W_n^φ alterek páronként merőlegesek egymásra – a (5.3) alatti skaláris szorzatra nézve – és ezek uniója is mindenütt sűrű $L^2(\mathbb{R})$ térben. A W_n^φ alterek, ψ_k^n $(k \in \mathbb{Z})$ alakú, bázisfüggvényeit *waveletekenek* nevezzük. A fejezet későbbi részeiben még foglalkozunk ezek szerkesztésével.

Most a Fourier-analízis eszközeit használva a Riesz-bázisok néhány fontos tulajdonságát mutatjuk be. Az ortonormált rendszerek mellett biortogonális rendszereket is vizsgálunk. Nevezetesen akkor mondjuk, hogy a $(\varphi_k, k \in \mathbb{Z})$ és $(\psi_k, k \in \mathbb{Z})$ $L^2(\mathbb{R})$ térbeli rendszerek *biortogonálisak*, ha

$$< \varphi_k, \psi_k >= \delta_{kl} \ , \ (k, l \in \mathbb{Z}). \tag{5.9}$$

Mint láttuk – a 2.6. szakaszban – a Fourier-transzformáció unitér, ezért a biortogonális rendszerek Fourier-transzformáltja is biortogonális. A

$$\varphi_k^n = 2^{n/2}\delta_{2^n}(\tau_k\varphi) = 2^{n/2}\tau_{k2^{-n}}(\delta_{2^n}\varphi), \quad (k, n \in \mathbb{Z})$$

függvények Fourier-transzformáltja (2.15) alapján

$$\mathcal{F}\varphi_k^n = 2^{n/2}\epsilon_{k2^{-n}}\mathcal{F}(\delta_{2^n}\varphi)$$

alakba írható. A most leírtak segítségével megmutatható, hogy az alábbi három állítás ekvivalens [8]:

1. $\forall n \in \mathbb{Z}$ esetén a $(\varphi_k^n, k \in \mathbb{Z})$ Riesz-bázis.

2. A $(\varphi_k^0, k \in \mathbb{Z})$ rendszer Riesz-bázis.

3. Léteznek olyan $0 < m \leq M < \infty$ állandók, melyekre

$$m \leq \sqrt{E_1(|\mathcal{F}\varphi|^2)} \leq M, \tag{5.10}$$

   ahol E_1 a 2.2. szakaszban bevezetett (1-szerint) periodizáló operátort jelöli.

Ez utóbbi – (5.10) – feltételnek elegettevő $\varphi \in L^2(\mathbb{R})$ fügvények halmazát jelöljük $\mathcal{R}$ szimbólummal. Az előbbi állítások ekvivalenciájának következménye, hogy az $\mathcal{R}$ halmazt azok a φ függvények alkotják, amelyekből (5.2) alapján képzett φ_k^n rendszerek minden n egész szám esetén Riesz-bázist alkotnak. A Fourier-transzformáció segítségével jellemezni lehet a (5.8) alatt bevezetett V_n^φ alterek egymáshoz való viszonyát. Nevezetesen a V_n^φ tér Fourier-transzformáltja előállítható

$$\widehat{V_n^\varphi} := \{\mathcal{F}f : f \in V_n^\varphi\} = \{\lambda\delta_{2^{-n}}(\mathcal{F}f) : \lambda \in L_{2^n}^2\} \tag{5.11}$$

alakban, továbbá a $(\tau_k\varphi, k \in \mathbb{Z})$ és $(\tau_k\psi, k \in \mathbb{Z})$ rendszerek akkor és csak akkor biortogonálisak, ha

$$E_1(\mathcal{F}\varphi\overline{\mathcal{F}\psi}) = 1. \tag{5.12}$$

Ez utóbbi állítás következménye, hogy a $(\tau_k\varphi, k \in \mathbb{Z})$ rendszer akkor és csak akkor ortonormált, ha

$$E_1(|\mathcal{F}\varphi|^2) = 1. \tag{5.13}$$

Végül a V_n^φ és V_n^ψ alterek akkor és csak akkor ortogonálisak, ha

$$E_1(\mathcal{F}\varphi\overline{\mathcal{F}\psi}) = 0. \tag{5.14}$$

Ezek után már nyilvánvalónak tűnik az a tény, hogy minden transzláció invariáns Riesz-bázis által generált altér ortonormált bázissal is generálható. Nevezetesen minden $\varphi \in \mathcal{R}$ esetén létezik olyan $\psi \in L^2(\mathbb{R})$, melyre $(\tau_k\psi, k \in \mathbb{Z})$ ortonormált rendszer és $V_n^\varphi = V_n^\psi$ egyenlőség minden n egész számra teljesül.

5.3. Skálázási egyenlet

Az MR felbontás $V_n^\varphi \subseteq V_{n+1}^\varphi$ monotonitási feltétele a (5.11) alapján azzal ekvivalens, hogy alkalmas $\lambda \in L^2_{2^{n+1}}(\mathbb{R})$ fügvénnyel fennáll a

$$\delta_{2^{-n}}(\mathcal{F}\varphi) = \lambda\delta_{2^{-n-1}}(\mathcal{F}\varphi)$$

egyenlőség. Vezessük be az $\alpha := \delta_{2^{n+1}}\lambda$ jelölést, majd alkalmazzuk mindkét oldalra a $\delta_{2^{n+1}}$ operátort, ekkor az alábbi, az eredetivel ekvivalens feltételhez jutunk:

$$(\mathcal{F}\varphi)(2x) = \alpha(x)(\mathcal{F}\varphi)(x), \quad \text{ahol} x \in \mathbb{R}, \alpha \in L^2_1. \tag{5.15}$$

Ez utóbbi (5.15) feltételt *skálázási egyenletnek* szokták nevezni. Az α függvényt a φ-hez tartozó *aluláteresztő (alulvágó) szűrőnek* nevezzük. A (2.16) azonosság és az $\alpha \in L^2_1$ szűrő $\alpha(x) = \sum_{k\in\mathbb{Z}} c_k\epsilon_k(x)$ $(c_k, k \in \mathbb{Z}) \in \ell^2$ Fourier-soros előállításának következményeként a (5.15) skálázási egyenlettel ekvivalens alakhoz jutunk:

$$\frac{1}{2}\varphi(\frac{1}{2}x) = \sum_{k\in\mathbb{Z}} c_k\varphi(x-k), \quad x \in \mathbb{R}, \tag{5.16}$$

ahol a c_k az α szűrőhöz tartozó trigonometrikus Fourier-együtthatók sorozata. Ez utóbbi egyenlőség mindkét oldalát 2-vel beszorozva, majd elvégezve az $x \to 2x$ helyettesítést, a következő – gyakorlatban is jól használható – formulához jutunk:

$$\varphi(x) = 2\sum_{k\in\mathbb{Z}} c_k\varphi(2x-k), \quad x \in \mathbb{R}. \tag{5.17}$$

Mivel a skála függvények általában kompakt tartójúak, ezért a fenti egyenlőségben az összegzés véges sok tagból áll. Jól láthatóan, a fenti egyenletben egyaránt szerepel a φ folytonos függvény és a c_k diszkrét együtthatók. A kulcs azonban az x és a $2x$ időhöz tartozó skála függvények egyidejű jelenléte. E nélkül ugyanis csak egy egyszerű konstans-együtthatós egyenletet kéne megoldanunk.

Példaként alkalmazzuk a fenti egyenletet a Haar esetre, ahol az együtthatók $c_0 = c_1 = 1/2$, a skála függvény pedig a $[0, 1)$ intervallum doboz (box) -függvénye. Ennek következményeképp $2c_0 = 1$ és $2c_1 = 1$, ami miatt a (5.17) egyenlőség a következő egyszerű alakot ölti:

$$\varphi(x) = \varphi(2x) + \varphi(2x - 1).$$

Tehát $\varphi(x)$ gráfját összenyomtuk 2-vel, hogy megkapjuk $\varphi(2x)$ gráfját. Amikor eltoljuk 1/2-del megkapjuk a $\varphi(2x - 1)$-hez tartozó gráfot. Az eredményből leolvasható, hogy a Haar-együtthatókhoz tartozó skála függvény a következő alakot ölti:

$$\varphi(x) = \begin{cases} 1 & \text{ha } 0 \le x < 1, \\ 0 & \text{egyébként,} \end{cases}$$

aminek gráfja nyilvánvalóan a doboz-függvény.

A továbbiakban jelöljük $\mathcal{M}$-mel azoknak a $\varphi \in \mathcal{R}$ függvényeknek az összességét, amelynek van nem-negatív, páros, a $[0, \infty)$ intervallumon monoton fogyó $L^1(\mathbb{R})$-beli majoránsa, továbbá eleget tesz a

$$\widehat{\varphi}(0) = (\mathcal{F}\varphi)(0) = \int_{-\infty}^{\infty} \varphi(t)\ dt = 1 \tag{5.18}$$

feltételnek. Ezek után írjuk fel a skálázási egyenletet x helyett $2^{-1}x, 2^{-2}x, \ldots,$ $2^{-n-1}x$ helyekre, ekkor a következő egyenlőséghez jutunk:

$$\widehat{\varphi}(x) = \alpha(\frac{1}{2}x)\widehat{\varphi}(\frac{1}{2}x) = \cdots = \alpha(\frac{1}{2}x)\alpha(\frac{1}{2^2}x)\cdots\alpha(\frac{1}{2^n}x)\widehat{\varphi}(\frac{1}{2^n}x). \tag{5.19}$$

Mivel $\varphi \in \mathcal{M}$ esetén $\widehat{\varphi}$ folytonos és $\widehat{\varphi}(0) = 1$, ezért innen $\alpha(0) = 1$ következik, továbbá $n \to \infty$ határátmenettel az MR-felbontást generáló függvény Fourier-transzformáltjának az alábbi szorzatelőállítását kapjuk:

$$\widehat{\varphi}(x) = \prod_{n=1}^{\infty} \alpha(\frac{1}{2^n}x) \quad (x \in \mathbb{R}). \tag{5.20}$$

Tehát megmutattuk, hogy a (5.15) skálázási egyenlet bámely $\mathcal{M}$-beli megoldásának Fourier-transzformáltja alkalmas $\alpha \in L^2_1, \alpha(0) = 1$ feltételnek elegettevő függvénnyel előállítható (5.20) alakú konvergens végtelen szorzatként. Megfordítva, ha valamely $\alpha \in L^2_1, \alpha(0) = 1$ szűrőből képzett

$$A(x) = \prod_{n=1}^{\infty} \alpha(\frac{1}{2^n}x) \quad (x \in \mathbb{R}). \tag{5.21}$$

végtelen szorzat majdnem mindenütt konvergens, akkor az A függvény nyilván majdnem minden $x \in \mathbb{R}$ pontban eleget tesz az $A(2x) = \alpha(x)A(x)$ egyenletnek és az $A(0) = 1$ feltételnek. Az eddigiekből következik, hogy ha valahogy garantálni tudjuk, hogy $A \in L^2(\mathbb{R})$, akkor létezik olyan $\varphi \in L^2(\mathbb{R})$, amelyre $\widehat{\varphi} = A$, tehát a (5.15) egyenlet valóban fennáll. Egyszerűen belátható, hogy ha az α 1-szerint periodikus, folytonos függvény a 0 pontban, $L > 0$ konstanssal eleget tesz az

$$|\alpha(x) - \alpha(0)| \leq L|x| \quad (x \in \mathbb{R}) \tag{5.22}$$

Lipschitz-féle feltételnek, akkor a (5.21) végtelen szorzat minden kompakt intervallumon egyenletesen konvergens, következésképpen az A függvény folytonos.

Most vizsgáljuk meg, hogy milyen α szűrők generálnak MR-felbontást. Bizonyítható, hogy ha $\varphi \in \mathcal{M}$ olyan függvény, melyre (5.13) teljesül, akkor fennáll a (5.15) skálázási egyenlet, továbbá az $\alpha \in L_1^2$ 1-szerint periodikus függvényre

$$\alpha(0) = 1, \quad |\alpha(x)|^2 + |\alpha(x + \frac{1}{2})|^2 = 1 \quad (x \in \mathbb{R}) \tag{5.23}$$

egyenlőségek igazak. Megfordítva, tegyük fel, hogy a folytonosan differenciálható, 1-szerint periodikus α szűrő eleget tesz a (5.23) feltételnek. Ekkor a (5.20) végtelen szorzat konvergens és

$$A(x) := \lim_{m\to\infty} \prod_{n=1}^{m} \alpha(\frac{1}{2^n}x) \tag{5.24}$$

határfüggvénye $L^2(\mathbb{R})$-beli. Továbbá, ha az

$$A_m := \chi_{\mathcal{I}_m} \prod_{n=1}^{m} \delta_{2^{-m}}\alpha \quad (\mathcal{I}_m := [-2^{m-1}, 2^{m-1}), m \in \mathbb{N}^*) \tag{5.25}$$

részsorozatoknak van egy közös $L^2(\mathbb{R})$-beli majoránsa, akkor a (5.24) alatt értelmezett A függvényre $E_1(|A|^2) = 1$ teljesül, melynek következtében a $\widehat{\varphi} = A$ alapján értelmezett φ függvény MR-felbontást generál. Az állítások következménye, hogy, ha α szűrő eleget tesz a (5.23) feltételnek és $|\alpha(x)| > 0$ teljesül minden olyan x-re, melyre $|x| \leq 1/4$, akkor a fent – (5.24) alatt – definiált $A \in L^2(\mathbb{R})$ függvényre $E_1(|A|^2) = 1$, következésképpen a $\widehat{\varphi} = A$ alapján értelmezett φ függvény a tér egy MR-felbontását generálja [8].

A skálázási egyenlet végtelen szorzatos előállítása $\widehat{\varphi}$-ra fennáll anélkül, hogy a φ függvényekre ortogonalitási feltételt szabnánk. Ez utóbbi megkövetelése azt jelentené, hogy a $\varphi(x - k)$ eltolt skála függvényekre ortogonalitást várunk el. A következőekben megmutatjuk, hogy ezen ortogonalitás miként hat a frekvencia-tartományra és látni fogjuk, hogy ez ekvivalens egy $C(x) \equiv 1$ alakú feltétellel, ahol C egy 1-szerint periodikus függvény. Mint ilyen függvény, létezik Fourier-soros $\sum c_k e^{i2\pi kx}$ előállítása. Vegyük a c_k együtthatók következő, skaláris szorzattal való előállítását:

$$c_k = 2\pi \int_{-\infty}^{\infty} \varphi(t)\overline{\varphi(t-k)}\, dt \tag{5.26}$$

Ennek következtében, a Parseval formula miatt $C(x)$ a következő alakba írható:

$$C(x) = \sum_{-\infty}^{\infty} |\widehat{\varphi}(x+n)|^2. \tag{5.27}$$

Ekkor az eltolt $\varphi(x-k)$ függvények ortogonalitása a $C(x) \equiv 1$ feltétellel ekvivalens. Ugyanis alkalmazva $\varphi(t)$ skálafüggvényre és annak $\varphi(t-k)$ eltoltjára a Fourier-transzformációt – az (2.15) azonosság figyelembe vételével – a kapott függvények $\widehat{\varphi}(x)$ és $e^{-i2\pi kx}\widehat{\varphi}(x)$ alakúak. Ekkor a (5.26) alatti skaláris szorzat a következő alakot ölti:

$$c_k = \int_{-\infty}^{\infty} \widehat{\varphi}(x)\overline{\widehat{\varphi}(x)}e^{i2\pi kx}\, dx = \int_0^1 \sum_{-\infty}^{\infty} |\widehat{\varphi}(x+n)|^2 e^{i2\pi kx}\, dx,$$

ami nem más, mint C függvény k-adik Fourier együtthatója, ezért $C(x) = \sum c_k e^{i2\pi kx}$. Egy FIR – azaz véges impulzus válaszú – szűrő [1] esetében $\varphi(t) = 0$ a kompakt $[0, N]$ intervallumon kívül, ami a függvény tartóját képezi. Ez által a (5.26) alatti skaláris szorzatban $c_k = 0$ minden olyan k-ra, melyre $|k| > N$, ugyanis ekkor a $\varphi(t)$ és $\varphi(t-k)$ skála függvényekhez tartozó gráfokban nincs átfedés.

Amikor az eltolt skála függvények ortogonálisak, akkor minden c_k skaláris szorzat nulla, kivéve a c_0-hoz tartozót, melyre $c_0 = 1$. Ennek következményeképp $\varphi(t-k)$ ortonormáltak akkor és csak akkor, ha a (5.27) formulára $C(x) \equiv 1$ teljesül. Ortonormált szűrőkészlet előállításához használjuk a (5.23) alatti α szűrőhöz tartozó egyenlőséget. Ugyanis $\alpha(x)$ segítségével megkapjuk $\widehat{\varphi}(x)$ függvényt, amiből előállíthatjuk $C(x)$-et. Amennyiben C folytonos, úgy a $C(x) \equiv 1$ feltétel ekvivalens a (5.23) alatti egyenlőséggel.

Azonban vannak olyan esetek amikor a fenti $C(x) \equiv 1$ nem, vagy csak nehezen garantálható. Most egy olyan módszert mutatunk, amellyel az említett feltétel mégi kielégíthető ezekre az esetekre. Nevezetesen, tegyük fel, hogy a $\varphi(t-k)$ skála függvények nem ortonormáltak. Osszuk le $\widehat{\varphi}(x)$ függvényt az adott $C(x)$ függvénnyel (sőt, még jobb, ha annak négyzetgyökével), hogy megkapjuk az új, már ortogonált $\widehat{\varphi}_{ort}(x)$ függvényt:

$$\widehat{\varphi}_{ort}(x) = \frac{\widehat{\varphi}(x)}{\sqrt{C(x)}}.$$

Ebből – az előzőek szerint – közvetlenül megkaphatjuk az új $\varphi_{ort}(t-k)$ bázist:

$$C_{ort}(x) = \sum \left| \frac{\widehat{\varphi}(x+k)}{\sqrt{C(x)}} \right|^2 = \frac{C(x)}{C(x)} \equiv 1,$$

ami teljesül, ha $C(x)$ sehol sem nulla. Ez utóbbi feltétel pedig azzal ekvivalens, hogy a szóban forgó rendszer Riesz bázis.

Most induljunk ki a skálázási egyenlet (5.16) alatti alakjából. Általában az alkalmazások azokat az MR-felbontásokat használják, amelyek φ generátor függvényei *kompakt tartójúak*. Belátjuk, hogy ebben az esetben a generátorhoz tartozó szűrők *trigonometrikus polinomok*. Tehát tegyük fel, hogy a φ generátor függvény kompakt tartójú, azaz

$$\left\{ x \in \mathbb{R} \middle| \varphi(x) \neq 0 \right\} \subseteq [a, a+N] \, ,$$

legyen továbbá $\varphi_k^0 = \tau_k \varphi$ $(k \in \mathbb{Z})$ egy ortonormált rendszer. Ekkor a (5.16) skálázási egyenlet c_k együtthatóira fennáll a

$$c_k = \frac{1}{2} \int_{-\infty}^{\infty} \varphi\left(\frac{t}{2}\right) \overline{\varphi(t+k)} \, dt \quad (k \in \mathbb{Z})$$

egyenlőség. Mivel

$$\begin{aligned} \varphi\left(\frac{t}{2}\right) &= 0 \quad (t \notin (2a, 2a+2N)), \\ \varphi(t+k) &= 0 \quad (t \notin (a-k, a-k+N)), \end{aligned}$$

ezért $c_k = 0$ minden olyan esetben, amikor $k \notin (-a-2N, -a+N)$, melynek következtében

$$\alpha = \sum_{k=-a-2N}^{-a+N} c_k \epsilon_k$$

valóban egy trigonometrikus polinom.

5.4. Ortonormált waveletek

Most szerkesszük meg az $L^2(\mathbb{R})$ térnek egy ortonormált wavelet bázisát, kiindulva valamely V_n^{φ} $(n \in \mathbb{Z})$ MR-felbontást generáló φ függvényből. Az eddigiek alapján már kitalálható, hogy olyan $\psi \in L^2(\mathbb{R})$ függvényt – anya-waveletet – keresünk $||\psi||_{L^2(\mathbb{R})} = 1$ feltétel mellett, hogy az általa generált

$$\psi_k^n(x) := 2^{n/2}\psi(2^n x - k), \quad x \in \mathbb{R}, k, n \in \mathbb{Z} \tag{5.28}$$

wavelet rendszer ortonormált legyen $L^2(\mathbb{R})$ téren, azaz teljesüljön a

$$< \psi_k^n, \psi_l^m >= \delta_{kl}\delta_{nm} \quad (k, l, n, m \in \mathbb{Z}) \tag{5.29}$$

egyenlőség. Itt jegyezzük meg, hogy az eddigi vizsgálatainkban, az MR-felbontások $(\varphi_k^n, k \in \mathbb{Z})$ rendszere csak rögzített n-re – más szóval szintenként – alkotott ortonormált rendszert, a különböző szinteken lévő függvényekre már nem teljesült az ortogonalitás. Azonban most olyan rendszert akarunk szerkeszteni, hogy annak bármely két különböző – akár más-más szinteken lévő – tagja ortogonális. A továbbiakban feltesszük, hogy mind φ, mind ψ eleget tesz a (5.13) feltételnek. Megmutatjuk, hogy ekkor létezik olyan β szűrő, hogy a $\delta_2(\widehat{\psi}) := \beta\widehat{\varphi}$ alapján értelmezett $\psi \in L^2(\mathbb{R})$ fügvénnyel fennáll a V_0^φ altér

$$V_0^\varphi = V_{-1}^\varphi \oplus W_{-1}^\psi \tag{5.30}$$

ortogonális felbontása. Ebből ugyanis – teljes indukcióval – az következik, hogy (5.30) teljesül 0 helyett tetszőleges $n \in \mathbb{Z}$ indexre, amelynek következtében a V_{n+1}^φ terek következő felbontásához jutunk:

$$V_{n+1}^\varphi = V_0^\varphi \oplus W_0^\psi \oplus W_1^\psi \oplus \cdots \oplus W_n^\psi. \tag{5.31}$$

Ebből már következik, hogy a (5.13) és (5.30) feltételeket kielégítő ψ függvényből képzett (5.28) rendszer ortonormált. A (5.31) felbontás következtében azonban gyakran a φ_k^0, ψ_k^n $(k \in \mathbb{Z}, n \in \mathbb{N})$ ortonormált rendszert használjuk.

A továbbiakhoz vezessünk be egy – skaláris szorzathoz nagyon hasonlító – jelölést:

$$[f, g] := f\overline{g} + \tau_{\frac{1}{2}}(f\overline{g}), \quad (f, g \in L_1^2), \tag{5.32}$$

ami nem más, mint L_1^2-beli függvénypárok, 1/2-szerint periodikus függvényeknek való megfeleltetése a skaláris szorzat szokásos tulajdonságainak megtartása mellett. Azonban megjegyzendő, hogy e szorzat értékei nem számok, hanem 1/2-szerint periodikus függvények. A tulajdonságokból következően definiálható két $\gamma_1, \gamma_2 \in L_1^2$ függvények []-*ortonormalitása* a

$$[\gamma_i, \gamma_j] = \delta_{i,j} \quad (i, j = 1, 2) \tag{5.33}$$

formula segítségével. Ekkor az $[f,\gamma_i]$ $(i=1,2)$ 1/2-szerint periodikus függvényeket az $f\in L_1^2$ függvény γ_i rendszer szerinti []-*Fourier-együtthatóinak* nevezzük. Belátható, hogy a

$$\gamma_1 := \frac{1}{\sqrt{2}},\quad \gamma_2 := \frac{\epsilon}{\sqrt{2}}$$

rendszer []-ortonormált, sőt, általában a $\gamma\in L_1^2$ []-normált függvényből képzett

$$\gamma_1 := \gamma,\quad \gamma_2 := \epsilon\tau_{\frac{1}{2}}(\overline{\gamma}) \tag{5.34}$$

függvénypár []-ortonormált. Továbbá bármely $\gamma_1,\gamma_2\in L_1^2$ []-ortonormált függvénypárral az L_1^2 tér minden f eleme előállítható

$$f = [f,\gamma_1]\gamma_1 + [f,\gamma_2]\gamma_2 \quad (f\in L_1^2) \tag{5.35}$$

[]-*Fourier-sor* alakjában. Legyen most $\gamma_1,\gamma_2\in L_1^2$, $c\in L^2(\mathbb{R})$ és tegyük fel, hogy a

$$\alpha(2x) = \gamma_1(x)c(x),\quad \beta(2x) = \gamma_2(x)c(x),\quad E_1(|c|^2) = 1 \tag{5.36}$$

egyenlőségek fennállnak. Ekkor igaz a következő azonosság:

$$\delta_2 E_1(\alpha\overline{\beta}) = [\gamma_1,\gamma_2]. \tag{5.37}$$

Ennek felhasználásával könnyen igazolható, hogy ha a (5.13) feltételt kielégítő φ és ψ függvényekkel fennáll a (5.30) felbontás, akkor léteznek olyan $\alpha,\beta\in L_1^2$ szűrők, hogy

$$\delta_2\widehat{\varphi} = \alpha\widehat{\varphi},\ \delta_2\widehat{\psi} = \beta\widehat{\psi}, \tag{5.38}$$

$$[\alpha,\alpha] = [\beta,\beta] = 1,\ [\alpha,\beta] = 0. \tag{5.39}$$

Megfordítva, ha $E_1(|\widehat{\varphi}|^2) = 1$ és fennállnak (5.38) és (5.39) feltételek, akkor a (5.38) alatt definiált ψ függvényre $E_1(|\widehat{\psi}|^2) = 1$, továbbá érvényes a (5.30) felbontás. Ezek szerint bármely V_n^φ $(n\in\mathbb{N})$ MR-felbontást generáló φ skála-függvény ismeretében (5.34) alapján megszerkeszthetjük egy teljes ortonormált wavelet rendszer anya-waveletét. Nevezetesen legyen $\varphi\in L^2(\mathbb{R})$ olyan függvény, melyre $E_1(|\widehat{\varphi}|^2) = 1$ és $\delta_2(\widehat{\varphi}) = \alpha\widehat{\varphi}$ teljesül. Ugyanakkor értelmezzük β és ψ függvényeket a következő képpen:

$$\beta := \epsilon\tau_{\frac{1}{2}}(\overline{\alpha}),\quad \delta_2\widehat{\psi} := \beta\widehat{\psi}. \tag{5.40}$$

Ekkor $(V_n^\varphi, n\in\mathbb{N})$ az $L^2(\mathbb{R})$ tér egy MR-felbontása, továbbá a $(\psi_k^n, k,n\in\mathbb{Z})$ egy teljes ortonormált rendszer [8]. A (5.40) alatt értelmezett $\beta\in L_1^2$ szűrő

Fourier-együtthatói kifejezhetők az α szűrő együtthatóival. Nevezetesen, ha α következő

$$\alpha = \sum_{k \in \mathbb{Z}} c_k \epsilon_k$$

Fourier-soros előállítását tekintjük, akkor az

$$\int_0^1 \beta(t)\overline{\epsilon_k}(t)\ dt = \int_0^1 \overline{\alpha}(t)\overline{\epsilon_{k-1}}(t)\ dt = \overline{\int_0^1 \alpha(t)\overline{\epsilon_{1-k}}(t)\ dt}$$

egyenlőségek következtében β Fourier-soros előállítása a

$$\beta = \sum_{k \in \mathbb{Z}} \overline{c_{1-k}} \epsilon_k$$

alakot ölti.

5.5. Folytonos Wavelet-Transzformáció

A wavelet-transzformáció elvégezhető folytonos időben (függvényeken) és diszkrét időben (vektorokon) egyaránt. A bemenet ennek megfelelően tehát vagy egy $f(x)$ függvény, vagy egy $x(n)$ vektor. A kimenet mindkét esetben a wavelet együtthatók egy halmaza, amelyek kifejezik az input jelet a wavelet bázisban. Függvények és végtelen hosszú jelek esetében természetesen ez a bázis szükségszerűen végtelen. A következőekben megadjuk a folytonos wavelet-transzformáció pontos definícióját, majd az inver-formula egzakt alakját.

A folytonos wavelet-transzformáció értelmezéséhez rögzítsünk egy $\psi \in C_0(\mathbb{R}), \psi \neq 0$ függvényt és az integráloperátor magfüggvényeire vezessük be a

$$\psi_s^t(x) := \frac{1}{\sqrt{|s|}} \overline{\psi}\left(\frac{x-t}{s}\right) \quad (x \in \mathbb{R}, (s,t) \in \mathbb{R}^* \times \mathbb{R}) \tag{5.41}$$

jelölést, ahol $\mathbb{R}^* = \mathbb{R} \backslash \{0\}$ és $C_0(\mathbb{R})$ az $\mathbb{R}$-en értelmezett, végtelenben eltűnő folytonos függvények halmaza. A fenti képletben természetesen s a skála értékét, mígy t az eltolást jelöli, amelyek ebben az esetben folytonos paraméterek. Ekkor az $f \in L^1(\mathbb{R})$ függvény *wavelet-transzformáltját (CWT)* a

$$(\mathcal{W}_\psi f)(s,t) := \int_{\mathbb{R}} f(x)\psi_s^t(x)\ dx \quad ((s,t) \in \mathbb{R}^* \times \mathbb{R}) \tag{5.42}$$

utasítással értelmezzük. A ψ függvényt *anya-waveletnek* nevezzük. Könnyen megmutatható, hogy a (5.42) leképezés egy korlátos, líneáris transzformáció az $L^1(\mathbb{R})$ téren, melyre

$$\sup_{(s,t)\in\mathbb{R}^*\times\mathbb{R}} \sqrt{|s|}\ |(\mathcal{W}_\psi f)(s,t)| \leq ||\psi||_\infty ||f||_1 \tag{5.43}$$

teljesül, továbbá $\mathcal{W}_\psi f \in C_0(\mathbb{R}^* \times \mathbb{R})$ [8].

A wavelet transzformáció esetében ψ-re – $\psi \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$, $\psi \neq 0$ mellett – még azt is kikötjük, hogy eleget tesz a

$$\int_{\mathbb{R}^s} |(\mathcal{F}\psi)(t)|^2 \frac{dt}{|t|} = 1 \tag{5.44}$$

normálási feltételnek. Ez a feltétel minden olyan ψ fügvény esetében teljesül, amelyre alkalmasan választott $a, b > 0$ és $C > 0$ számokkal fennállnak a

$$\widehat{\psi}(0) = \int_{\mathbb{R}} \psi(s)\ ds = 0,$$

$$|\widehat{\psi}(t)| \leq C|t|^a \quad (|t| \leq b), \quad |\widehat{\psi}(t)| \leq C|t|^{-a} \quad (|t| > b)$$

egyenlőtlenségek.

Az operátor $L^2(\mathbb{R})$ térre való kiterjesztéséhez vezessük be a következő $L^2_\lambda(\mathbb{R}^* \times \mathbb{R})$ téren értelmezett skaláris szorzatot:

$$\ll F, G \gg := \int_{\mathbb{R}^*} \int_{\mathbb{R}} F(s,t)\overline{G(s,t)} \frac{ds\ dt}{s^2} \quad (F, G \in L^2_\lambda(\mathbb{R}^* \times \mathbb{R})), \tag{5.45}$$

ahol λ a fenti téren értelmezett $d\lambda(s,t) := \frac{ds\ dt}{s^2}$, $(s \in \mathbb{R}^*, t \in \mathbb{R})$ mértéket jelöli. Ezek után tegyük fel, hogy a ψ függvény kielégíti a (5.43) normálási feltételt. Ekkor a kiterjesztett $\mathcal{W}_\psi : L^2(\mathbb{R}) \to L^2_\lambda(\mathbb{R}^2)$ transzformáció esetén bármely $f, g \in L^2(\mathbb{R})$ fügvénypárra

$$\ll \mathcal{W}_\psi f, \mathcal{W}_\psi g \gg = < f, g >, \tag{5.46}$$

tehát a kiterjesztett operátor *unitér* [8].

Ezen kívül megmutatható, hogy a wavelet transzformáció *injektív*. Tegyük fel, hogy ψ kielégíti a $\psi \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$, $\psi \neq 0$ feltételeket és $f \in L^1(\mathbb{R}) \cup L^2(\mathbb{R})$. Ha $\mathcal{W}_\psi f = 0$, akkor $f = 0$ az $\mathbb{R}$ majdnem minden pontjában. Ismeretes, hogy ha $\psi \in L^1(\mathbb{R})$ és $f \in L^1(\mathbb{R}) \cup L^2(\mathbb{R})$ esetén a $\mathcal{W}_\psi f$ második változó szerinti Fourier-transzformáltja a következő egyszerű alakban írható fel:

$$\mathcal{F}((\mathcal{W}_\psi f)(s,\cdot)) = \sqrt{|s|}\mathcal{F}f \cdot \delta_s(\mathcal{F}\psi_-) \quad (s \in \mathbb{R}^*), \tag{5.47}$$

ahol $\psi_-(x) = \overline{\psi}(-x)$. Ezek után nem nehéz belátni, hogy $\mathcal{W}_\psi$ inverze kifejezhető a Fourier-transzformáció inverzével. Tegyük fel ugyanis, hogy

$\psi \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ kielégíti a (5.43) normálási feltételt, továbbá legyen $f \in L^1(\mathbb{R})$ és $F := \mathcal{W}_\psi f$. Ha $\widehat{f} \in L^1(\mathbb{R})$, akkor majdnem minden $x \in \mathbb{R}$ pontban

$$f(x) = \int_{\mathbb{R}^*} \left(\int_{\mathbb{R}} F(s,t) \overline{\psi_s^t}(x) \; dt \right) \frac{ds}{|s|^2}. \tag{5.48}$$

A skála értékek és az eltolás paraméterek diszkretizálásával a fenti (5.48) inverziós formula a következő egyszerűbb alakot ölti:

$$f(x) = \sum_{n,k} b_{nk} \psi_n^k(x), \tag{5.49}$$

ahol b_{nk} súlyok a $(\mathcal{W}_\psi f)(n,k)$ wavelet együtthatókat jelölik.

6. fejezet

Szűrők és szűrőkészletek

6.1. Szűrők

A szűrő egy líneáris, idő-invariáns operátor, aminek bemenete egy vektor, jelöljük ezt a tömböt x-szel. A kimenete szintént egy – y-nal jelölt – tömb, ami nem más, mint az x-nek egy rögzített h vektorral való (3.12) konvolúciója:

$$y(n) = \sum_{k \in \mathbb{Z}} h(k)x(n-k). \tag{6.1}$$

Ez utóbi h tömb tartalmazza a szűrő együtthatóit, azaz elemei rendre a $h(0), h(1), \ldots$ értékekkel azonosíthatók. Amennyiben a szűrő digitális (diszkrét) – tehát nem analóg (folytonos) – úgy ezek az együtthatók diszkrét mintavételezésből kaphatók meg $t = nT$ időben, ahol T a mintavételezési periódus. Az $x = (\ldots, 0, 1, 0, \ldots)$ bemenetnek – az úgynevezett *egységimpulzusnak* – különlegesen nagy jelentősége van a jelfeldolgozásban. Ugyanis ezen bemenettel vett konvolúció egyetlen nem nulla értéket szolgáltat a kimenetén, ami nem más, mint $h(n)$. Ez az $y(n) = h(n)$ kimenet az n időponthoz tartozó válasza az egységimpulzusnak. Az így kapott $h(0), h(1), \ldots, h(N)$ sorozatot impulzus válasznak nevezzük.

Mint láthattuk, a konvolúció – a (2.19) azonosság szerint – a frekvenciatartományban a transzformált sorozatok szorzatába megy át, így lehetőség nyílik a konvolúció, gyors algoritmusokkal történő kiszámítására (lásd 4.4. szakasz). Ez által nem meglepő, hogy az egész jelfeldolgozás a szűrők idő és frekvencia-tartománybeli viselkedésére épült. A fejezet további részeiben az $x(n)$ sorozat diszkrét Fourier-transzformációjára és z-transzformációjára a következő jelöléseket használjuk:

$$X(e^{i\omega}) = \sum_{n=-\infty}^{\infty} x(n)e^{-i\omega n} \quad \text{(DFT jelfeldolgozásbeli alakja)}, \tag{6.2}$$

$$X(\omega) = \sum_{n=-\infty}^{\infty} x(n)e^{-i\omega n} \quad \text{(DFT egyszerűsített alakja)}, \tag{6.3}$$

$$X(z) = \sum_{n=-\infty}^{\infty} x(n)z^{-n} \quad \text{(z-transzformáció alakja)}, \tag{6.4}$$

ahol ω egy frekvencia-tartománybeli pont, z pedig egy teszőleges komplex szám. Mint láthattuk, ez utóbbi formulába z helyére beírva az $e^{i\omega}$ egységkör egy pontját, visszakapjuk a Fourier-transzformált fogalmát.

Tekintsük bemenetnek az egységimpluzus 1-gyel késleltetett impulzusát. Ekkor – a szűrő línearitása és idő-invarianciája miatt – a kimenet is 1-gyel késleletve lesz. Azonban ez a késleltetés a frekvencia-tartományban már $e^{-i\omega}$ impulzusnak felel meg a (2.15) alatti transzláció (azaz késleltetés) és Fourier-transzformáció közötti kapcsolat következményeként. Ez által a frekvencia-tartoménybeli kimenet $Y(\omega) = e^{-i\omega}H(\omega)$ alakú lesz, ami nem más tehát, mint a késleltetett $h(n-1)$ impulzus válasza.

Természetesen a (5.3). szakaszban bevezetett, skála függvényekre vonatkozó (5.17) skálázási egyenlethez hasonló formula a waveletek W_n terén is bevezethető. A (5.4). szakaszban láthattuk, hogy a waveletek tere nem más, mint a skála függvények terének ortogonális kiegészítője, tehát tartalmazza a skála függvények két „szomszédos" skálájú tere közötti részleteket (különbségeket). Azt is láttuk, hogy a W_n tér ψ_n^k waveletei (bázis függvényei) nem csak ezen téren ortogonálisak egymásra, hanem minden más skálázáshoz tartozó W_j terek függvényeire is azok. Ezen bázis függvényeket háromféleképpen is meg lehet határozni:

1. W_n térből,
2. ψ anyawaveletből és
3. d_k wavelet-együtthatókból.

Az elsőt akkor használjuk, ha adottak a V_n skála terek, melyek különbségéből meghatározható W_n. A második használatával direkt módon – skálázással és eltolással – határozhatjuk meg wavelet bázisainkat. A harmadikat akkor célszerű alkalmazni, amikor a c_k szűrő-együtthatók adottak, ekkor az úgynevezett *alternating flip* [2] módszert alkalmazva meghatározhatók a d_k együtthatók a következő képpen: $d_k = (-1)^k c_{N-k}$.

Alkalmazva a Haar esetre a harmadik módszert, ahol az anyawavelet a jól ismert (5.4) alakot ölti. Ekkor $c_0 = c_1 = 1/2$, melynek következtében $d_0 = 1/2$ és $d_1 = -1/2$. A kapott együtthatókból felépített *wavelet egyenlet* a $\psi(x) = \psi(2x) - \psi(2x-1)$ alakot ölti, ami általánosabb formában a következőképpen fest:

$$\psi(x) = 2\sum_{k\in\mathbb{Z}} d_k\psi(2x-k) \quad x \in \mathbb{R}. \tag{6.5}$$

A (6.5) egyenlőség formálisan láthatóan megyegyezik a (5.17) alatti skálázási egyenlettel. A különbség, hogy most waveleteket állítunk elő, nem skála függvényeket, tehát az egyenlet két különböző skálájú wavelet teret kapcsol össze. Ábrázolva a fenti Haar esethez tartozó egyenlet gráfját, a $[0, 1)$ intervallum jól ismert lépcsős függvénye köszön vissza.

A skálázási egyenlet megahatározásakor megadtuk az *aluláteresztő szűrő (lowpass filter, LP)* pontos matematikai deinícióját. Most ezeket vizsgáljuk meg egy picit más szemszögből a sepciális Haar példán keresztül. Ez a legegyszerűbb szűrő-konstrukció, amelynek kimenete az $x = n$ és $x = n-1$ időkhöz tartozó bemenetek (lokális) átlaga:

$$y(n) = \frac{1}{2}x(n) + \frac{1}{2}x(n-1).$$

Látható, hogy a szűrő-együtthatók értéke $h_0 = h_1 = \frac{1}{2}$, ami azt jelenti, hogy a (6.1) konvolúciós összeg csupán a fenti két – $k = 0$ és $k = 1$ indexekhez tartozó – tagra redukálódik. Ezt a szűrőt a szakirodalomban szokás *mozgó átlagnak* [2] is nevezni, hiszen a kimenet az aktuális $x(n)$ komponens és az ezt megelőző átlaga. Ez által a speciális $x = (\dots, 0, 0, 1, 0, 0, \dots)$ egységimpulzusra adott kimenet $y = (\dots, 0, 0, \frac{1}{2}, \frac{1}{2}, 0, \dots)$ alakú lesz.

Szemléletesen, a líneáris, idő-invariáns LP szűrő a következő operátorok segítségével írható fel:

$$\text{átlag szűrő} = \frac{1}{2}(\text{identikus leképezés}) + \frac{1}{2}(\text{késleltetés}).$$

Ismeretes, hogy minden líneáris leképezés megadható mátrix formában. Mivel a bemeneti és kimeneti jelek végtelen hosszúak is lehetnek, azért a transzformációt megvalósító – 6.1 ábrán látható – H mátrix is szükségképpen az. Az ábrán látható H mátrix diagonálisa jelöli az „identikus leképezést”, míg az átló alatti elemek a „késleltetést” szimbolizálják. A mátrix alsóháromszög tulajdonsága azt a tényt tükrözi, hogy az eltaláunk használt szűrő *kauzális* (lásd a z-transzformációról szóló fejezet), tehát a kimenet nem keletkezhet előbb, mint a bemenet. A példánkban csak véges számú (2) nem nulla szűrőegyütthatója van $h(n)$-nek, ami azt jelenti, hogy a szűrő *véges impulzus válaszú (FIR)*.

$$\begin{bmatrix} \cdot \\ y(-1) \\ y(0) \\ y(1) \\ \cdot \end{bmatrix} = \begin{bmatrix} \cdot & 0 & \cdots & \cdots & 0 \\ \frac{1}{2} & \frac{1}{2} & \ddots & & \vdots \\ 0 & \frac{1}{2} & \frac{1}{2} & \ddots & \vdots \\ \vdots & \ddots & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \cdots & 0 & \cdot & \cdot \end{bmatrix} \cdot \begin{bmatrix} \cdot \\ x(-1) \\ x(0) \\ x(1) \\ \cdot \end{bmatrix}$$

6.1. ábra. Az $y = h * x$ konvolúció $y = Hx$ mátrix LP szűrőre vonatkozó alakja

Most vizsgáljuk meg a frekvencia-tartomány felől az aluláteresztő szűrőt. Ehhez – a szemléletesség kedvéért most – válasszuk a bemenetet a tiszta $x(n) = e^{i\omega n}$ frekvenciák sorozatára. Ennek egy speciális tulajdonsága, hogy a kimenet a bemenet – ω-től függő – számszorosa. A mi mozgó átlagunkra ez a következőket jelenti:

$$y(n) = \frac{1}{2}x(n) + \frac{1}{2}x(n-1) = \frac{1}{2}e^{in\omega} + \frac{1}{2}e^{i(n-1)\omega} = \left(\frac{1}{2} + \frac{1}{2}e^{-i\omega}\right)e^{in\omega}.$$

Az utóbbi zárójeles kifejezés nem más, mint a *frekvencia válasz függvény* $H(\omega) = \frac{1}{2} + \frac{1}{2}e^{-i\omega}$, amiből leolvasható a szűrő aluláteresztő volta, hiszen $\omega = 0$ közelében a $H(\omega)$ válasz 1 közeli értéket vesz fel. A kapott függvények álatlában komplex értékűek, azonban egy egyszerű gráffal ezek nem igazán jellemezhetőek. Ezért válasszuk szét a függvényt amplitúdóra $|H(\omega)|$ és fázisra $\phi(\omega)$, azaz:

$$H(\omega) = |H(\omega)|e^{i\phi(\omega)}.$$

A mi példánkra alkalmazva:

$$H(\omega) = \frac{1}{2} + \frac{1}{2}e^{-i\omega} = \frac{1}{2}\left(e^{i\omega/2} + e^{-i\omega/2}\right) = \left(\cos\frac{\omega}{2}\right)e^{-i\omega/2},$$

melyből már leolvasható az amplitúdó $(\cos\frac{\omega}{2})$ és a fázis $(-\frac{\omega}{2})$. A cos függvény tulajdonságaiból is látszik, hogy az valóban egy aluláteresztő szűrő, hiszen az $\omega = \pm\pi$-ben az amplitúdó értéke 0, míg az $\omega = 0$ pontban 1.

A fázis alakjáról pedig leolvasható, hogy a példa *líneáris fázisú*, amely tulajdonságnak csak néhány speciális szűrő felel meg. Ez tükrözi, hogy az együtthatók szimmetrikusak $(1/2, 1/2)$, azaz megfordítva az együtthatók sorozatát az eredmény nem változik. Sőt, ha ez a szimmetria a $n = 0$ pont körül teljesül, akkor a kapott frekvencia válaszok *valósak*. Azonban ez utóbbi tulajdonság nem állhat fent kauzális szűrők esetében, ugyanis az $n < 0$ értékek azonosan nullának tekintendők (természetesen a szimmetria fennálhat, de nem a 0 pont körül). Tehát egy líneáris fázisú szűrő lehet

$$\begin{bmatrix} \cdot \\ y(-1) \\ y(0) \\ y(1) \\ \cdot \end{bmatrix} = \begin{bmatrix} \cdot & 0 & \cdots & \cdots & 0 \\ -\frac{1}{2} & \frac{1}{2} & \ddots & & \vdots \\ 0 & -\frac{1}{2} & \frac{1}{2} & \ddots & \vdots \\ \vdots & \ddots & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \cdots & 0 & \cdot & \cdot \end{bmatrix} \cdot \begin{bmatrix} \cdot \\ x(-1) \\ x(0) \\ x(1) \\ \cdot \end{bmatrix}$$

6.2. ábra. Az $y = h * x$ konvolúció $y = Hx$ mátrix HP szűrőre vonatkozó alakja

1. $h(k) = h(N - k)$ (szimmetrikus), vagy
2. $h(k) = -h(N - k)$ (antiszimmetrikus).

Mint láttuk az alulátersztő szűrő a bemenő jelek lokális átlagát veszi, mintegy kisimítva az egyenetlen részeket. Ugyanis az ilyen emelkedések magas frekvenciájú komponensek, amelyeket az alulátersztő szűrő kivág. Ezzel szemben a *felülvágó szűrő (highpass filter, HP)* a jelek lokális különbségén operál, ami viszont a sima részeket nm engedi át. Ugyanis a sima részek alacsony frekvenciájú komponenseket takarnak. A szemléletesség kedvéért ezen szűrőket is a speciális Haar példán keresztül mutatjuk be.

A felüláteresztő szűrő felfogható az aluláteresztő szűrő tükrének, ugyanis ez a *mozgó különbségnek* [2] is titulált rendszer a bemenő jel (lokális) különbségeit veszi:

$$y(n) = \frac{1}{2}x(n) - \frac{1}{2}x(n-1).$$

Az egyenlőségből is leolvasható, hogy az új szűrő együtthatók $h_0 = \frac{1}{2}$ és $h_1 = -\frac{1}{2}$. Ez által a (6.1) alatti konvolúció ismételten egy két tagú összegbe megy át, mert a felüláteresztő szűrő vektora $h = (\ldots, 0, 0, \frac{1}{2}, -\frac{1}{2}, 0, 0, \ldots)$ alakú. Ugyanis h éppen az egységimpulzusra adott válaszfüggvénye, azaz $y = h$.

Szemléletesen a HP szűrőt a következő operátorok segítségével adhatjuk meg:

$$\text{felüláteresztő szűrő} = \frac{1}{2}(\text{identikus leképezés}) - \frac{1}{2}(\text{késleltetés}).$$

Mint minden líneáris transzformáció, természetesen ennek is van mátrix alakja, amelyet a 6.2. ábrán szemléltetünk. Ez a FIR szűrő szintén egy kauzális (alsóháromszög alakú) szűrő, azonban a frekvenciaválasza teljesen különbözik a $H_0(\omega) = \frac{1}{2}(1 + e^{-i\omega})$ aluláteresztő szűrőjétől.

A frekvencia-tartománybeli szemléltetéshez, a bemenetnek tekintsük ismét az $x(n) = e^{in\omega}$ tiszta frekvenciák sorozatát. Ekkor a kimenet:

$$y(n) = \frac{1}{2}x(n) - \frac{1}{2}x(n-1) = \frac{1}{2}e^{in\omega} - \frac{1}{2}e^{i(n-1)\omega} = \left(\frac{1}{2} - \frac{1}{2}e^{-i\omega}\right)e^{in\omega}.$$

alakú, ahol az utolsó zárójel nem más, mint a felüláteresztő szűrő válasz függvénye, azaz $H_1(\omega) = \frac{1}{2}(1 - e^{-i\omega})$. A kapott frekvencia választ bontsuk fel ismét amlitúdóra és fázisra, kiemelve $H_1(\omega)$-ból $e^{-i\omega/2}$-t:

$$H_1(\omega) = \frac{1}{2} - \frac{1}{2}e^{-i\omega} = \frac{1}{2}\left(e^{i\omega/2} - e^{-i\omega/2}\right) = \left(\sin\frac{\omega}{2}\right)ie^{-i\omega/2}.$$

Az eredményből leolvasható, hogy az amplitúdó most $|h_1(\omega)| = |\sin\frac{\omega}{2}|$, ahol az abszolút jel a szinusz függvény $[-\pi, 0)$ intervallumon való negatív volta miatt nem hagyható el. A kapott eredményből leolvasható, hogy a szűrőnk valóban felüláteresztő, hiszen az $\omega = \pm\pi$-hez közeli pontokhoz tartozó értékek 1 közeliek, tehát ezeket a jeleket átengedi, míg az $\omega = 0$ ponthoz torlódó értékeket nem.

A $H_1(\omega)$ fázis tényezője egy picit trükkösebb a komplex i jelenléte miatt:

$$e^{i\phi(\omega)} = \begin{cases} -ie^{-i\omega/2} & \text{ha } -\pi < \omega < 0, \\ +ie^{-i\omega/2} & \text{ha } 0 < \omega < \pi. \end{cases}$$

A 0 ponthoz tartozó ugrás miatt a fázis nem folytonos, azonban – e tény fölött szemethúnyva – erre is azt mondjuk, hogy *líneáris fázisú.*

A jelfeldolgozásban nagyon fontos tulajdonság a *rekonstruálhatóság*, azaz, hogy mikor, milyen feltételek mellett lehet a feldolgozott jelet visszaállítani annak eredeti voltába. Ennek szükséges feltétele, hogy az alkalmazott transzformációk invertálhatóak legyenek, vagyis a megvalósitó mátrixnak létezzen az inverze. Azonban az eddig bemutatott alul-, illetve felüláteresztő szűrők mátrixai nem invertálhatóak. Ugyanis egy H leképezés invertálható, és $Hx = 0$, akkor $H^{-1}Hx = 0$, ami miatt $x = 0$. Ugyanakkor a frekvencia válasza egy invertálható leképezésnek nem lehet nulla sehol sem, azaz $H(\omega) \neq 0$ egyetlen ωra sem teljesülhet, azonban a mi szűrőink ezt nem teljesítik, ugyanis $H_0(\pi) = 0$ és $H_1(0) = 0$. Azonban e két szűrőt egyszerre alkalmazva, a kapott *szűrőkészlet* nemcsak, hogy invertálható, de még szét is választja a bemenő jeleket frekvencia sávokra. Azonban – mivel minkét szűrőt alkalmaztuk – a kimeneten kétszer annyi jel fog megjelenni. Ezen hatás kiküszöbölésére vezették be az úgynevezett *downsampling*, illetve az inverz transzformáció elvégzéséhez szükséges *upsampling* műveleteket [2]. A fenti operációkkal részletesebben a következő 6.2. szakaszban foglalkozunk, míg a szűrőkészleteket a 6.3. részben tárgyaljuk.

6.2. Downsampling és Upsampling

Szűrőkészlet tervezéséhez az alapgondolat, hogy osszuk fel a bejövő jeleket különböző frekvencia sávokra. A szemléletesség kedvéért, mi most két sávra szeparálunk egy alulteresztő C és egy felüláteresztő D szűrő segítségével. Mint ahogyan azt a (6.3). szakaszban látni fogjuk, a szűrőkészlet transzformációi elkülöníthetőek *analízis* és *szintézis* folyamatokra. Ezek közül az első foglalkozik az adott jel frekvencia sávokra bontásával, míg utóbbi a feldolgozott jelet transzformálja vissza az eredeti, kiindulási állapotába. Amennyiben ez a visszaállított jel egy az egyben megegyezik az eredeti jellel, úgy azt mondjuk, hogy a transzformációk sorozata egy *tökéletes rekonstrukciót* hajtott végre. Az ilyen típusú szűrőkészleteket a legfontosabbak a jelfeldolgozásban, úgyhogy a tervezésnél mi is erre törekszünk.

Tehát egy szűrőkészletben alkalmazva a C és D szűrőinket, elkerülhetetlen az a hatás, hogy a kimeneten kétszer annyi jel jelenjen meg, mint amennyi a bemeneten volt. Ennek kiküszöbölésére egy nagyon szép megoldást vezettek be, melynek neve: *downsampling*. Az angol elnevezés arra utal, hogy a sáv szűrés kimenetéből csak bizonyos elemeket tartsunk meg. Mivel most csak két szűrőt alkalmazunk, ezért ebben az esetben a Cx és Dx outputjainak csak a *felét tartjuk meg*. $(\downarrow 2)$ szimbólummal jelölve a downsampling műveletét, megadjuk annak adott jelre kifejtett hatását:

$$(\downarrow 2)\begin{bmatrix} \cdot \\ y(-1) \\ y(0) \\ y(1) \\ y(2) \\ \cdot \end{bmatrix} = \begin{bmatrix} \cdot \\ y(-2) \\ y(0) \\ y(2) \\ y(4) \\ \cdot \end{bmatrix}$$

Látható tehát, hogy tulajdonképpen nem történet más, mint bemenő jelsorozatból kiválogattuk a páros indexűeket, míg a többi elemet elhagytuk. Ennek köszönhetően az adott jel nagyságát felére csökkentettük.

Azonban, egyszerű meggondolások alapján észrevehető, hogy a dowsampling művelete nem invertálható, azaz a páratlan indexű tagok elvesznek az operáció elvégzése során. Példaként tekintsük az

$$x_1(n) = (-1)^n = (\ldots, 1, -1, 1, -1, \ldots)$$

bemenetet, majd alkalmazzuk rá a downsamplin műveletét kapjuk:

$$v(n) = (\downarrow 2)(\ldots, 1, -1, 1, -1, \ldots) = (\ldots, 1, 1, 1, 1, \ldots).$$

Látható, hogy ugyanazt a kimenetet kapjuk, mintha az

$$x_2(n) = (\ldots, 1, 1, 1, 1, \ldots)$$

bemeneti vektorra alkalmaztuk volna a downsamplinget. Érdemes egy kicsit jobban megvizsgálni ezt a két bemenetet. Az $x_1(n)$-hez tartozó bemenet az $\omega = \pi$ frekvenciához tartozik, ugyanis $x_1(n) = e^{i\pi n} = (-1)^n$. Ezzel szemben a csupa egyesből álló $x_2(n)$ bemenet az $\omega = 0$ frekvenciából állítható elő, ugyanis $x_2(n) = e^{i0n} = 1$. Ezt a jelenséget – amikor két bemeneti jelsorozathoz tartozó kimenet megegyezik – a szakirodalomban *aliasing*-nak nevezik [2].

Észrevehető azonban, hogy ha megkötést teszünk a bemeneti jel frekvenciájára, akkor ez a jelenség megszüntethető. Nevezetesen határoljuk sávokra a bemeneti jel frekvenciáját felső, vagy alsó sávokra a következő képpen:

$$X(\omega) = 0 \text{ ha } 0 \le |\omega| < \frac{\pi}{2}, \text{ vagy } X(\omega) = 0 \text{ ha } \frac{\pi}{2} \le |\omega| < \pi.$$

Ugyanis, ha tudjuk, hogy a bemeneti ω frekvencia biztosan kisebb, mint $\frac{\pi}{2}$, akkor az $x_2(n)$ bemenet nem lehet a kiinduló jelsorozat, ami által mégis invertálni tudjuk a downsampling műveletét. És – mint ahogyan azt az előző 6.1. szakaszban láthattuk – a C, illetve D, alul-, illetve felüláteresztő szűrők pontosan ezt a sávokra bontást végzik el, maguk után vonva a downsamplingre vonatkozó invertálhatóság feltételének teljesülését. Éppen ezért olyan fontos – és egyben elengedhetetlen – e művelet előtti *szűrés*.

Egy jobb szűrőkészlet két különböző C és D szűrőt használ. A tervezés része, annak biztosítása, hogy a $(\downarrow 2)Cx$ és $(\downarrow 2)Dx$ tartalmazza mindazt az információt, ami a bemenet visszaállításhoz kell. A downsampling műveletének formális definíciója a következő:

$$\text{ha } v = (\downarrow 2)x, \text{ akkor } v(n) = x(2n). \tag{6.6}$$

Természetesen ez a leképezés is megadható mátrix formában a kövekező képpen:

$$\begin{bmatrix} \cdot & 1 & & & & & \\ & 0 & 0 & 1 & & & \\ & & & 0 & 0 & 1 & \\ & & & & & \cdot & \cdot \end{bmatrix} \cdot \begin{bmatrix} \cdot \\ x(-2) \\ x(-1) \\ x(0) \\ x(1) \\ x(2) \\ \cdot \end{bmatrix} = \begin{bmatrix} \cdot \\ x(-2) \\ x(0) \\ x(2) \\ \cdot \end{bmatrix}$$

Látható, hogy a mátrix megkapható az egységmátrixból, a páratlan indexű sorok elhagyásával. A $(\downarrow 2)$ tulajdonképpen egy líneáris operátor. Mivel a mátrix sorai ortonormáltak, ezért a mátrix $(\downarrow 2)^T$ transzponáltja a leképezés egyoldali inverze (a mátrix végtelen dimenziója miatt csak egyoldali).

A most bevezetett $(\downarrow 2)^T$ operátor nagyon fontos, ugyanis a *downsampling* *transzponáltja az upsampling* művelet:

$$(\downarrow 2)^T = (\uparrow 2).$$

Az operáció nem csinál mást, mint nullákat rak a páratlan indexű helyekre, szemléletesen:

$$(\uparrow 2)\begin{bmatrix} \cdot \\ v(-1) \\ v(0) \\ v(1) \\ \cdot \end{bmatrix} = \begin{bmatrix} \cdot \\ v(-1) \\ 0 \\ v(0) \\ 0 \\ v(1) \\ 0 \\ \cdot \end{bmatrix}$$

Formálisan definiálva az $u = (\uparrow 2)v$ upsampling műveletét a következő

$$u(n) = \begin{cases} v(k) & \text{ha } n = 2k, \\ 0 & \text{ha } n = 2k+1 \end{cases} \tag{6.7}$$

formulát kapjuk.

Természetesen ez a transzformáció is megadható mátrix alakban:

$$\begin{bmatrix} \cdot & & & & \\ 1 & 0 & & & \\ & 0 & & & \\ & 1 & 0 & & \\ & & 0 & & \\ & & 1 & 0 & \\ & & & \cdot & \end{bmatrix} \cdot \begin{bmatrix} \cdot \\ v(-1) \\ v(0) \\ v(1) \\ \cdot \end{bmatrix} = \begin{bmatrix} \cdot \\ v(-1) \\ 0 \\ v(0) \\ 0 \\ v(1) \\ \cdot \end{bmatrix},$$

amiből látható, hogy ez éppen a downsampling mátrixának transzponáltja. Továbbá az is észrevehető, hogy a két mátrixot összeszorozva $(\downarrow 2)(\uparrow 2)$ sorrendben, éppen az egységmátrixot kapjuk, azaz $(\downarrow 2)(\uparrow 2)x = x$. Azonban a szokott sorrend éppen fordított, azaz először a downsampling műveletét alkalmazzuk és aztán az upsamplinget. Ebben az eseben azonban $(\uparrow 2)(\downarrow 2) \neq I$.

Most vizsgáljuk meg a bevezetett műveleteket a frekvencia-tartományban. Ehhez válasszuk bemenetnek a tiszta $x(k) = e^{i\omega k}$ frekvenciát. Ekkor a k-adik komponense $v = (\downarrow 2)x$ vektornak $v(k) = e^{i2k\omega}$, ami szintén tisztán exponenciális, $\omega = 2\pi$ frekvenciával. Tehát a művelet megduplázta a frekvenciát.

Elsőre úgy tűnhet, hogy $V(\omega) = X(\frac{\omega}{2})$. Azonban alkalmazva a downsampling műveletét az $x'(k) = e^{1(\omega+\pi)k}$ bemenő jelekre, az $e^{i2k\pi} = 1$ azonosság miatt $v(k) = v'(k)$ egyenlőséghez jutunk. Kaptuk tehát, hogy a két kimenet megegyezik, azaz a 2ω frekvencia megjelenik, ha megduplázzuk ω-t, vagy, ha megdoplázzuk $\omega+\pi$-t. Formálisan tehát a $v = (\downarrow 2)x$ downsampling hatása a frekvencia-tartományban a következő:

$$V(\omega) = \frac{1}{2}\left[X\left(\frac{\omega}{2}\right) + X\left(\frac{\omega}{2} + \pi\right)\right]. \tag{6.8}$$

Ugyanis legyen $u = (\ldots, x(0), 0, x(2), 0, \ldots)$, ahol a nem nulla értékek a páros indexű elemek. A definició következménye, hogy $(\downarrow 2)u = (\downarrow 2)x$. Véve ennek Fourier-transzformációját:

$$U(\omega) = \sum_{n|2} x(n)e^{-in\omega} = \frac{1}{2}\sum_{n} x(n)e^{-in\omega} + \frac{1}{2}\sum_{n} x(n)e^{-in(\omega+\pi)},$$

ahol az utóbbi tag második összegzésében a páratlan n-nek kiejtik egymást, ugyanis $e^{-in\pi} = (-1)^n$. A fenti formulát egyszerűbb alakra hozva:

$$U(\omega) = \frac{1}{2}\left[X(\omega) + X(\omega + \pi)\right].$$

Ennek segítségével már könnyen meghatározhatjuk $V(\omega)$-t, véve a frekvencia felét, azaz

$$V(\omega) = U\left(\frac{\omega}{2}\right),$$

ami a (6.8) alakkal ekvivalens.

Mint azt az időtartoményi elemzésnél láttuk, a nem sávszűrt jelnél megjelenik az aliasing. Ez most abban a formában mutatkozhat, hogy az $X\left(\frac{\omega}{2}\right)$-höz és az $X\left(\frac{\omega}{2} + \pi\right)$-hoz tartozó gráf tartalmaz átfedést. Természetesen megfelelően sávszűrt jeleknél ez az átlapolódás nem jelentkezik, tehát az $X\left(\frac{\omega}{2} + \pi\right)$ aliasért felelős tag kiszűrhető.

Az idő-tartományban az upsampling műveletét két részre kellett bontani a (6.7) formulának megfelelően. A frekvencia-tartományban ennél egyszerűbb a formális definíció. Ugyanis felírva a Fourier-transzformációt, csak a páros indexű tagok jelennek meg, ugyanis $u(2k + 1) = 0$, minden k-ra:

$$U(\omega) = \sum u(n)e^{-in\omega} = \sum u(2k)e^{-i2k\omega} = \sum v(k)e^{-i2k\omega} = V(2\omega).$$

Tehát az $u = (\uparrow 2)v$ upsampling, frekvencia-tartománybeli hatása formálisan az

$$U(\omega) = V(2\omega). \tag{6.9}$$

egyenlőséggel egyezik meg.

Mivel $V(\omega)$ 2π szerint periodikus, ezért $V(2\omega)$ már csak π szerint lesz periodikus, tehát az eredeti gráf besűrűsödik $|\omega| \leq \pi$-ről az $|\omega| \leq \frac{\pi}{2}$ intervallumra. Ez által egy új kép jelenik meg éppen a gráf mellet, e jelenség neve *imaging* [2]. Az imaging jelenség éppen az aliasing ellentéte, ugyanis ez utóbbi két különböző – ω és $\omega + \pi$ – frekvencia bemenetből ugyanazt a kimenetet produkálja, míg az imaging egy input ω frekvenciához két kimeneti – $\frac{\omega}{2}$ és $\frac{\omega}{2} + \pi$ – frekvenciát gyárt.

A downsampling művelet az analízis részhez, míg az upsampling művelet a szintézis folyamatához tartozik. Éppen ezért elég kézenfekvő megvizsgálni őket egyszerre, azaz

$$v = (\downarrow 2)x, \text{ majd } u = (\uparrow 2)v, \text{ hogy megkapjuk } u = (\uparrow 2)(\downarrow 2)x\text{-et.}$$

A transzformációja $u = (\uparrow 2)(\downarrow 2)x$ műveletsorozatnak

$$U(\omega) = \frac{1}{2}\left[X(\omega) + X(\omega + \pi)\right], \tag{6.10}$$

ami a downsampling és upsampling műveletek (6.8) és (6.9) alatti definícióinak következménye. Tehát egyszerre alkalmazva a két műveletet számolni kell mind az aliasing, mind az imaging jelenséggel.

Vizsgáljuk most meg a műveleteinket a z-tartományon, azaz a komplex számsíkon. Itt természetesen a Fourier-transzformáció helyett annak – 3.6. szakaszban bevezetett – általánosítását, a z-transzformációt használjuk. Ismeretes, hogy az összefüggés a két transzformáció között a $z = e^{i\omega}$ helyettesítéssel adható meg. Azonban a z-tartományon az ω, $\omega + 2\pi$, $\omega + 4\pi$, stb frekvenciákat lehetetlen megkülönböztetni, ugyanis a $z = e^{i\omega}$, $z = e^{i(\omega+2\pi)}$ és $z = e^{i(\omega+4\pi)}$ komplex számok ugyanazokat az értékeket jelölik.

A $z = e^{i\omega}$ egyenlőségnek megfelelően a következő azonosságok érvényesek:

1. a frekvencia felezése $\frac{\omega}{2}$-re a z-tartományban $z \to z^{1/2}$ helyettesítéshez vezet,
2. a frekvencia duplázása 2ω-ra a z-tartományban $z \to z^2$ helyettesítéshez vezet,
3. a frekvencia π-vel való eltolása ((3.33) alapján) a z-tartományban $z \to -z$ helyettesítéshez vezet.

Ezek után a $v = (\downarrow 2)x$ és az $u = (\uparrow 2)v$ downsampling és upsampling műveletek hatása a z-tartományban

$$V(z) = \frac{1}{2}\left[X(z^{\frac{1}{2}}) + X(-z^{\frac{1}{2}})\right], \text{ illetve } U(z) = V(z^2). \tag{6.11}$$

Ennek következménye, hogy az $u = (\uparrow 2)(\downarrow 2)x$ műveletsorozat transzformációja

$$U(z) = \frac{1}{2}(X(z) + X(-z))$$

alakú.

6.3. Szűrőkészletek

A szűrőkészlet nem más, mint szűrők egy halmaza. Ezen szűrőket is csoportosíthatjuk aszerint, hogy az adott jel feldolgozásáért, vagy visszaállításáért felelősek. Ebben a felfogásban tehát beszélhetünk *analízis* készletről és *szintézis* készletről [2]. Az analízis készlet gyakran két szűrőből áll, ezek a – már jól ismert – aluláterszt**ő**, illeve felüláteresztő szűrők. Láthattuk, hogy ezek a transzformációk különböző frekvencia sávokra bontják a bemenő jeleket. A megszűrt jelek mennyiségének decimálására – az előző 6.2. szakaszban – bevezettük a downsampling műveletét, melynek alkalmazásával elkerülhetjük a teljes kimenetnek a tárolását. Arra is rávilágítottunk, hogy miért fontos e művelet előtti szűrés.

Ha az 6.1. szakaszban bevezett H_0 és H_1 szűrőkre alkalmazzuk a downsampling eljárást – az output megfelezése miatt – energia vesztés keletkezik. Ennek kompenzálására szorozzuk meg a szűrőinket $\sqrt{2}$-vel:

$$C(\omega) = \sqrt{2}H_0(\omega), \text{ illetve } D(\omega) = \sqrt{2}H_1(\omega),$$

melyhez tartozó együtthatók

$$c_0 = c_1 = \frac{\sqrt{2}}{2} = \left(\frac{1}{\sqrt{2}}\right),$$

illetve

$$d_0 = \frac{\sqrt{2}}{2} = \left(\frac{1}{\sqrt{2}}\right) \text{ és } d_1 = -\frac{\sqrt{2}}{2} = \left(-\frac{1}{\sqrt{2}}\right).$$

A továbbiakban a $\sqrt{2}H_0$ és $\sqrt{2}H_1$ szűrők helyett, az egyszerűség kedvéért a C, illetve D jelölést használjuk.

Azonban a aluláteresztő – C-vel való – szűrést követő downsampling összevonható egyetlen L mátrixba a következő képpen:

$$L = (\downarrow 2)C = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & & & \\ & & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \\ & & & & \ddots \end{bmatrix}.$$

Ehhez hasonló módon, természetesen a felüláteresztő szűrőhöz tartozó decimálás is elvégezhető egyetlen műveletként:

$$B = (\downarrow 2)D = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & & & \\ & & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \\ & & & & \ddots \end{bmatrix}.$$

Látható, hogy egyik mátrix sem lesz diagonális, egy sor az előzőből duplaeltolással kapható meg. Az így kapott L és B szűrőkből má felépíthető az analízis mátrix:

$$\begin{bmatrix} (\downarrow 2)C \\ (\downarrow 2)D \end{bmatrix} = \begin{bmatrix} L \\ B \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & & & \\ & & 1 & 1 & \\ & & & & \ddots \\ -1 & 1 & & & \\ & & -1 & 1 & \\ & & & & \ddots \end{bmatrix}. \qquad (6.12)$$

A fenti mátrix teljes analízis készletet reprezentálja, ugyanis a vele végzett transzformáció az aluláteresztő és a felüláteresztő szűrés mindegyikét végrehajtja az adott jelen. A $\sqrt{2}$-vel való leosztás miatt a mátrix minden sora egységvektor (ezért a $\sqrt{2}$-vel való szorzás), melyek ortogonálisak egymásra.

Ugyanakkor az oszlopai is ortonormáltak, ezért a mátrix inverze a transzponáltja:

$$\begin{bmatrix} L \\ B \end{bmatrix}^{-1} = \begin{bmatrix} L^T & B^T \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & & -1 & & \\ 1 & & 1 & & \\ & 1 & & -1 & \\ & 1 & & -1 & \\ & \vdots & & & \vdots \end{bmatrix}. \qquad (6.13)$$

Ez utóbbi (6.13) mátrix adja a szintézis készletet. Látható, hogy az egy *ortogonális szűrő készlet*, hiszen az inverz éppen a transzponálttal egyenlő. Ez azt jelenti, hogy egyiket a másik után alkalmazva *tökéletes rekonstrukciót* érhetünk el. Azonban vannak olyan szűrőkészletek is, amikor a szintézis mátrix „csak" inverze az analízis mátrixnak. Ezeket *biortogonális szűrőkészletek* nevezzük [2]. Ebben az esetben a szintézis mátrixban szereplő együtthatók különböznek az analízis mátrix c_k, illetve d_k együtthatóitól.

Ortonormált szűrők esetén létezik egy nagyon egyszerű módszer a d_k együtthatók c_k együtthatókból történő előállítására. Az inverz transzformáció aluláteresztő szűrőjének együtthatóit f_k-val, míg a felüláteresztő szűrő együtthatóit g_k-val jelölve a következő fontos kapcsolatok állapíthatók meg:

$$\begin{aligned} d_k &= (-1)^k c_{N-k} && \text{(alternating flip)}, \\ f_k &= c_{N-k} && \text{(flip)}, \\ g_k &= (-1)^{k+1} c_k && \text{(alternating sign)}, \end{aligned}$$

ahol $0 \leq k \leq N$. Biortogonális esetben egy picit más a helyzet, ugyanis a felüláteresztő szűrők együtthatói függenek az aluláteresztő szűrők együtthatóitól:

$$\begin{aligned} d_k &= (-1)^k f_k, \\ g_k &= (-1)^{k+1} c_k. \end{aligned}$$

Alkalmazva egy x bemenő jelre a már bevezetett L és B aluláteresztő és felüláteresztő szűrőket az eredményt a következő szemléletes formában írhatjuk fel:

$$Lx = \frac{1}{\sqrt{2}} \begin{bmatrix} & \cdot & \\ x(0) & + & x(-1) \\ x(2) & + & x(1) \\ x(4) & + & x(3) \\ & \cdot & \end{bmatrix} \quad Bx = \frac{1}{\sqrt{2}} \begin{bmatrix} & \cdot & \\ x(0) & - & x(-1) \\ x(2) & - & x(1) \\ x(4) & - & x(3) \\ & \cdot & \end{bmatrix}. \tag{6.14}$$

A páratlan indexű sorokat mind az L, mind a B eltűntette, tehát eredményül két feleakkora kimenetet kaptunk. A rekonstrukció során a következő két vektort szeretnénk kapni:

$$w_0 = \frac{1}{2} \begin{bmatrix} & \cdot & \\ x(0) & + & x(-1) \\ x(0) & + & x(-1) \\ x(2) & + & x(1) \\ x(2) & + & x(1) \\ & \cdot & \end{bmatrix} \quad w_1 = \frac{1}{2} \begin{bmatrix} & \cdot & \\ -x(0) & + & x(-1) \\ x(0) & - & x(-1) \\ -x(2) & + & x(1) \\ x(2) & - & x(1) \\ & \cdot & \end{bmatrix}. \tag{6.15}$$

Ugyanis összeadva a két vektort, azaz elvégezve a $w_0 + w_1$ műveletet a kívánt vektort kapjuk vissza. Azonban a bemenet $x(n)$, míg a kimenet $x(n-1)$, azaz késleltetett. A késleltetés oka, hogy a szűrőink *kauzálisak*, ami miatt a kimenet nem keletkezhet előbb, mint a bemenet.

Egy jól felépített szintézis készlet az analízis készlet inverze. Mint láttuk az analízis folyamata két lépésből áll, szűrés, majd downsampling. A szintézis menete is két lépésből tevődik össze, nevezetesen az upsampling műveletét egy újabb szűrés követ.

Az első lépés ugyanis a szintézis során, hogy visszaállítsuk a jelet eredeti méretére. Ezt végzi el az upsampling művelete, ami a páratlan indexű helyeket feltölti nullákkal, mintegy kiterjesztve a jelet. Láthattuk, hogy az upsampling nem más, mint a downsampling transzponáltja. A vektorok, amelyeket az upsampling alkalmazása után kapunk (6.14) eredményekre alkalmazva:

$$u_0 = \frac{1}{\sqrt{2}} \begin{bmatrix} & . & \\ x(0) & + & x(-1) \\ & 0 & \\ x(2) & + & x(1) \\ & 0 & \\ x(4) & + & x(3) \\ & . & \end{bmatrix} \quad u_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} & . & \\ x(0) & - & x(-1) \\ & 0 & \\ x(2) & - & x(1) \\ & 0 & \\ x(4) & - & x(3) \\ & . & \end{bmatrix}. \tag{6.16}$$

A második lépés a szintézis során a szűrés, melynek során az (6.15) alatti eredményeket szeretnénk megkapni. A szűrés bemenete az upsampling kimenetén kapott u_0 és u_1 vektorok. A szintézis folyamatához kellő F és G szűrőinket az analízis készlet C, illetve D szűrőiből szeretnénk meghatározni. A mi esetünkben, amikor ismerjük a kívánt kimenetet, a két szűrő meghatározása egyszerű. Tehát F-fel akarjuk szűrni u_0-lát, hogy megkapjuk w_0-lát. Ha a bemenet egy komponense $u(n)$, akkor a kimenet komponenese

$$Fu(n) = \frac{1}{\sqrt{2}}(u(n) + u(n-1))$$

alakú lesz, amiből leolvasható, hogy a szűrő-együtthatók $f_0 = f_1 = \frac{1}{\sqrt{2}}$ lesznek. A G szűrőnk egy különbség szűrő lesz, amit u_1-re akarunk alkalmazni és w_1-et akarjuk megkapni vele:

$$Gu(n) = \frac{1}{\sqrt{2}}(-u(n) + u(n-1)).$$

G együtthatói tehát $g_0 = \frac{1}{\sqrt{2}}$ és $g_1 = -\frac{1}{\sqrt{2}}$.

Megfigyelhető, hogy a példánkban szereplő négy szűrőnk mindegyikében szerepel az $\frac{1}{\sqrt{2}}$ szorzó tényező. Egy kevésbé szimmetrikus esetben lehetne $\frac{1}{2}$ az egyik készletben és 1 a másikban. Észrevehető, hogy amit így kapnánk az tulajdonképpen nem lenne más, mint egy két pontos diszkrét Fourier-transzformáció.

Most a tökéletes rekonstrukció kérdését vizsgáljuk meg egy picit közelebbről. A szűrőkészletünk legyen most is két-csatornás, azaz álljon alul-, illetve felüláteresztő szűrőkből. Az analízis készlet szűrőit jelöljük H_0-lal, illetve H_1-gyel, míg a szintézis szűrőit F_0-lal, valamint F_1-gyel (a 0-lás index az aluláteresztő, míg az 1-gyes index a felüláteresztő szűrőt jelölik). Tegyük fel azt is, hogy a H_0 és H_1 frekvencia válasz függvényei tartalmaznak átfedést (azaz van aliasing jelenség), valamint, hogy sem az amplitúdó, sem a fázis nem *torzításmentes* [2]. Az F_0 és F_1 szintézis szűrőknek alkalmazkodniuk kell a H_0 és H_1 analízis szűrőkhöz abban az értelemben, hogy e rossz jelenségeket megszűntessék.

A tökéletes rekonstrukció egy nagyon fontos tulajdonság. A downsampling és upsampling műveletek nélkül, a rekonstrukció késés nélkül a $F_0H_0 + F_1H_1 = I$ formulával lenne ekvivalens. l lépést késve azonban a z-tartománybeli hatás a következőt jelentené:

$$F_0(z)H_0(z) + F_1(z)H_1(z) = z^{-l}. \tag{6.17}$$

Az aliasing jelenség felbukkanását onnan vehetjük észre, hogy a z-t tartalmazó komponens mellet megjelenik a $-z$-t tartalmazó is (ez ekvivalens a frekvencia-tartománybeli ω, $\omega + \pi$ egyidejű megjelenésével). Alkalmazva a sampling műveleteket megfigyelhető, hogy a $H_0(z)X(z)$ szorzatnak csak a páros hatványai maradnak meg:

$$(\uparrow 2)(\downarrow 2)H_0x \text{ transzformációja } \frac{1}{2}(H_0(z)X(z) + H_0(-z)X(-z)).$$

A páratlan komponensek kiesése miatt a kapott függvény páros. A szintézis során, az aliasingért felelős második tag, $F_0(z)$-vel megszorzódik. Ennek az aliasnak törölnie kell a másik – felülvágó szűrőt tartalmazó – csatornán lévő $F_1(z)H_1(-z)X(-z)$ aliasing tagot. Az aliasing jelenség törlésének feltétele tehát a következő:

$$F_0(z)H_0(-z) + F_1(z)H_1(-z) = 0. \tag{6.18}$$

A sampling operátorokat alkalmazva a (6.17) azonosság egy picit módosul, ugyanis a jobb oldalra bejön még egy plussz 2-tes szorzó. Ez által a torzításmentesség feltétele, l késéssel a

$$F_0(z)H_0(z) + F_1(z)H_1(z) = 2z^{-l} \tag{6.19}$$

alakba írható.

Ahhoz, hogy a 2-csatornás szűrőkészletünk tökéletes rekonstrukciót hajtson végre (más szóval a szűrőkészletünk biortogonális legyen) mindkét – (6.18) és (6.19) alatti – feltételnek teljesülnie kell. A tökéletes rekonstrukció feltétele megadható mátrix formában is a következő képpen:

$$\begin{bmatrix} F_0(z) & F_1(z) \end{bmatrix} \begin{bmatrix} H_0(z) & H_0(-z) \\ H_1(z) & H_1(-z) \end{bmatrix} = \begin{bmatrix} 2z^{-l} & 0 \end{bmatrix}. \tag{6.20}$$

Tehát van most négy szűrőnk, amiket meg kell tervezni, figyelembe véve a fenti feltételeket. Elég nyilvánvalónak tűnik az az ötlet, hogy definiáljuk az egyik szűrőt a másikból a következő képpen:

$$F_0(z) := H_1(-z) \text{ és } F_1(z) := -H_0(-z). \tag{6.21}$$

Ezzel a választással az aliasing törlésének (6.18) feltétele automatikusan teljesül. A torzításmentesség (6.19) alatti egyenlőségét alakítsuk át egy kicsit. Nevezetesen definiáljuk az úgynevezett *szorzat szűrőt* [2] a következő képpen:

$$P_0(z) := F_0(z)H_0(z). \tag{6.22}$$

A fenti szűrő természetesen aluláteresztő. Definiáljuk a $P_1(z)$ felüláteresztő szűrőt azzel analóg módon az $F_1(z)$ és $H_1(z)$ szűrők szorzataként. Ha a szűrőinket a (6.21) módon definiáljuk, akkor fontos kapcsolatot vehetünk észre $P_0(z)$ és $P_1(z)$ között. Nevezetesen a

$$P_1(z) = -P_0(-z) \tag{6.23}$$

azonosság igaz. Ennek alkalmazásával a (6.19) torzításmentesség feltétele a

$$P_0(z) - P_0(-z) = 2z^{-l} \tag{6.24}$$

alakba írható.

Tehát egy 2-csatornás teljes rekonstrukciót megvalósító szűrőkészlet tervezése két lépésből áll:

1. lépés: *Tervezzük meg az aluláteresztő P_0 szűrőt, amely kielégíti a (6.24) feltételt.*

2. lépés: *Bontsuk P_0-lát F_0 és H_0 szorzatára, majd alkalmazzuk (6.21) egyenlőségeket, hogy meghatározhassuk F_1-gyet és H_1-gyet.*

A korábbiakban már szó volt a skálázási (5.17), illetve wavelet (6.5) egyenletekről, melyek – ha elvégezzük az energiavesztés kiegyenlítése miatti $\sqrt{2}$-vel való szorzást – a következő képpen módosulnak:

$$\varphi(x) = \sum \sqrt{2} c_n \varphi(2x - n), \tag{6.25}$$

illetve

$$\psi(x) = \sum \sqrt{2} d_n \psi(2x - n). \tag{6.26}$$

Ezen fontos egyenleteknek köszönhetően kapcsolatba tudjuk hozni egymással a waveleteket és a szűrőket. Ugyanis az aluláteresztő szűrő $c_0, \ldots, c_N$ együtthatói határozzák meg a $\varphi(x)$ skála függvényt, míg a felüláteresztő szűrő $d_0, \ldots, d_N$ együtthatói a $\psi(x)$ waveleteket állítják elő. Tehát három fő feladatunk van:

1. Kiszámolni az együtthatókat az

$$f_n(x) = \sum_k a_{nk}\varphi_n^k(x) \text{ és } f(x) = \sum_n \sum_k b_{jk}\psi_n^k(x)$$

   egyenletekből.

2. Megszerkeszteni $\varphi(x)$-et, megoldva a skálázási egyenletet.

3. Összekapcsolni $\varphi(x)$ és $\psi(x)$ tulajdonságait c és d szűrő-együtthatók tulajdonságaival.

A feladatok megoldását rekurzív módon végezzük el, mely közben látni fogjuk, hogy a függvények multirezolúciója miként kapcsolódik vektorok sávszűrésével. Tehát a három megoldás, amit picit jobban szemügyre veszünk a következő:

1. a_{nk} és b_{nk} meghatározása $a_{n+1,k}$-ból (és fordítva).

2. Egy rekurziós eljárás felállítása (Cascade algoritmus) $\varphi(x)$ megszerkesztésére.

3. $\varphi_n^k(x)$ és $\psi_n^k(x)$ ortogonalitásának bebizonyítása c és d ortogonalitásából.

Tegyük fel, hogy $f_1(x) \in V_1$. Ez nem más, mint a $\varphi_1^k(x) = \sqrt{2}\varphi(2x-k)$ bázis függvények lineáris kombinációja az 1. szinten. Mivel a multirezolúció a V_1 teret felbontja $V_1 = V_0 \times W_0$ terekre, ezért az $f_1(x)$ egyúttal a V_0 és W_0 terek kombinációja is, melyek bázis függvényei $\varphi_0^k(x) = \varphi(x-k)$, illetve $\psi_0^k(x) = \psi(x-k)$:

$$\sum a_{1k}\varphi_1^k(x) = \sum a_{0k}\varphi_0^k(x) + \sum b_{0k}\psi_0^k(x) = \sum a_{0k}\varphi(x-k) + \sum b_{0k}\psi(x-k). \tag{6.27}$$

Kiszámolunk egy bázis cserét. Adott a_{1k} V_1 bázisbeli skála függvény-együtthatókból meg akarjuk határozni a $V_0 \times W_0$ bázis a_{0k} skála-, illetve b_{0k} wavelet-együtthatóit. És ezt az eljárást folytatva minden n szintre, a V_{n+1} bázis $a_{n+1,k}$ együtthatóiból meg szeretnénk határozni a V_n, illetve W_n bázisok a_{nk}, illetve b_{nk} együtthatóit.

Feltételezve bázisaink ortonormalitását – hogy megtaláljuk a rekurziós megoldást – helyettesítsünk a (6.25) és (6.26) egyenletekbe n helyére $\ell - 2k$-t, kapjuk:

$$\varphi(x-k) = \sum \sqrt{2}c_n\varphi(2x-2k-n) = \sum \sqrt{2}c_{\ell-2k}\varphi_1^\ell(x), \tag{6.28}$$

illetve

$$\psi(x-k) = \sum \sqrt{2} d_n \psi(2x - 2k - n) = \sum \sqrt{2} d_{\ell-2k} \psi_1^\ell(x). \tag{6.29}$$

Szorozzuk meg mindkét egyenlet, mindkét oldalát $f_1(x)$-szel és integráljuk őket x szerint. A bázis függvények ortonormalitása miatt az integrál éppen a keresett a_{0k} és b_{0k} együtthatókat kapjuk vissza:

$$a_{0k} = \sum c_{\ell-2k} a_{1\ell} \ , \tag{6.30}$$

illetve

$$b_{0k} = \sum d_{\ell-2k} a_{1\ell} \ . \tag{6.31}$$

És ez a kulcsa a rekurziós megoldásnak. Ez éppen a szűrőkészlet hatása, amelynek bemenete az $a_{1\ell}$ együtthatók, kimenete pedig az a_{0k} és b_{0k} együtthatók. Az indexezésben észrevehetünk valami szokatlant, ugyanis egy átlagos konvolúció (downsamplig művelettel) $\sum c(2k-\ell) a_{1\ell}$ alakú lenne. A (6.30) rekurzióban megfigyelhető egy idő-fordulat a C szűrő és annak C^T transzponáltja között:

$$c_n^T = c_{-n} \text{ és } d_n^T = d_{-n}.$$

Tehát a multirezolúció egyik szintjéről, másik szintjére való áttérése, nem más, mint egy C^T, illetve D^T szűrővel való sávszűrés. Ugyanis a $V_{n+1} = V_n \times W_n$ téren értelmezett $\sum a_{n+1,\ell} \varphi_{n+1}^k(x)$ függvény együtthatói a_{nk} és b_{nk} az új ortonormált $\{\varphi_n^k, \psi_n^k\}$ bázisban:

$$a_{nk} = \sum_\ell c_{\ell-2k} a_{n+1,\ell} \ , \tag{6.32}$$

illetve

$$b_{nk} = \sum_\ell d_{\ell-2k} a_{n+1,\ell} \ . \tag{6.33}$$

A fenti egyenletek vektor alakja

$$a_n = (\downarrow 2) C^T a_{n+1} \ ,$$

illetve

$$b_n = (\downarrow 2) D^T a_{n+1} \ .$$

Az inverz formulához hajtsunk végre egy bázis csarét $\{\varphi_n^k, \psi_n^k\}$ bázisról, $\{\varphi_{n+1}^\ell\}$ bázisra. Az ortonormáltság következményeként az inverz műveletet a transzponált adja. Tehát az $a_{n+1,\ell}$ együtthatókat megkaphatjuk az a_{nk} és b_{nk} együtthatókból a szintézis szűrőkészlet alkalmazásával:

$$a_{n+1,\ell} = \sum_\ell c_{2k-\ell} a_{nk} + d_{2k-\ell} b_{nk} \ . \tag{6.34}$$

A $\varphi(x)$ skála függvény rekurziós megszerkesztéséhez induljunk ki a skálázási egyenlet 6.25 alakjából. A cél, hogy a bemenetként megadott $c_0, \ldots, c_N$ együtthatók segítségével előállítsuk $\varphi(x)$ függvényt. A megoldást az úgynevezett *Cascade algoritmus* adja.

Induljunk ki a $[0, 1]$ intervallum $\varphi^{(0)}(x)$ doboz-függvényéből, majd végezzünk iterációs lépéseket az aluláteresztő szűrővel a következőek szerint:

$$\varphi^{(i+1)}(x) = \sum_n \sqrt{2} c_n \varphi^{(i)}(2x - n) = \sum_n 2h_n \varphi^{(i)}(2x - n). \qquad (6.35)$$

Az algoritmus folytonos függvények felett operál. Ezek a függvények szakaszonként konstans értékűek, úgynevezett *lépcsős függvények*. Az iteráció előrehaladtával ezen szakaszok hosszai csökkennek (2^{-1}-vel). Amennyiben $\varphi^{(i)}(x)$ konvergál $\varphi(x)$ határ skála függvényhez, úgy a kapott függvény megoldása a skálázási egyenletnek.

A két különböző x és $2x$ skála egyidejű jelenléte, a downsampling folytonos változatának köszönhető. Ugyanis $(\downarrow 2)\varphi(n) = \varphi(2n)$ helyett most $(\downarrow 2)\varphi(x) = \varphi(2x)$ szerepel. A Cascade algoritmus valójában a szűrő mátrixon végzi az – akár végtelen sok – iterációs lépéseit, azaz $M = (\downarrow 2)2H$.

Az ilyen típusú $f(x)$ függvényeknek könnyen meg tudjuk feleltetni az $a(n)$ diszkrét analogonját. Ugyanis az i-edik intevallumon a függvény értéke éppen $a(i)$, más szóval $f(x) = a(i)$, ha $i \leq x < i + 1$.

Az egyszerűség kedvéért most tekintsünk el az iteráció egyes lépéseiben történő újaskálázástól és hajtsuk végre csak a h_k együtthatókkal végzett szűrést a z-tartományban, más szóval egy z-transzformációt:

$$H(z) = \sum_k h_k z^{-k}.$$

Ez tulajdonképpen a $H^{(1)}(z)$, első iterációs lépésnek felel meg. Most hajtsunk végre egy iterációs lépést, kapjuk

$$H^{(2)}(z) = H(z^2)H(z).$$

Ezzel analóg módon $H^{(3)}(z)$ a

$$H^{(3)}(z) = H(z^4)H(z^2)H(z)$$

formulával kapható meg. Az egyes iterációs lépések során az intervallumok finomodása miatt – az újraskálázás során – azok magassága megnő. Folytatva az iterációt a z-tartományban az i-edik lépés után a következő alakhoz jutunk:

$$H^{(i)}(z) = \prod_{k=0}^{i-1} H(z^{2^k}).$$

Tehát a Cascade algoritmus a z-tartományban nem más, mint a H aluláteresztő szűrővel való iteráció.

A 5.3. szakaszban a (5.18) - (5.25) alatti formulák segítségével már megmutattuk, hogy ez az eljárás mikor, milyen feltételek mellett lesz konvergens. Például a $h_0 = \frac{2}{3}$ és $h_1 = \frac{1}{3}$ együtthatókkal értelmezett szűrővel végzett iteráció sorozat nem fog konvergálni, hiszen a kapott $H(z) = \frac{2}{3} + \frac{1}{3}z^{-1}$ frekvencia válasz értéke a $z = -1$ pontban (tehát az $\omega = \pi$ helyen) nem nulla. Azonban ha a szűrő-együtthatókat a klasszikus Haar-együtthatóknak válasszuk – ahol $h_0 = h_1 = \frac{1}{2}$ – akkor az eljárás előállítja a keresett skála függvényt, ami nem lesz más, mint a kiinduló $\varphi^{(0)}(x)$ doboz-függvény.

Amikor egy szűrő készlet diszkrét időben ortogonális, akkor elvárható lenne, hogy ez a tulajdonság megmaradjon folytonos esetben is. Tehát minden $\psi(2^n x)$ wavelettől elvárjuk, hogy azok ortogonálisak legyenek a $\varphi(x-k)$ alakú skála függvényekre. Ezen kívűl a $\psi(2^n x - k)$ alakú waveletek kölcsönös ortogonalitását és a $\varphi(x-k)$ skála függvények kölcsönös ortogonalitását is megköveteljük. Természetesen a különböző szinteken lévő $\varphi(2x)$ és $\varphi(x)$ skála függvények ortogonalitása – a már látottak alapján – nem teljesülhet.

A továbbiakhoz tegyük fel, hogy a Cascade algoritmus egyenletesen konvergál x-ben, azaz $\varphi^{(i)}(x) \to \varphi(x)$. Amennyiben a c_k és d_k együtthatók egy ortonormált dzűrőkészletből valóak, úgy

1. A $\varphi(x-k)$ skála függvények ortonormáltak egymásra nézve:
$$\int_{-\infty}^{\infty} \varphi(x-n)\varphi(x-m) = \delta_{nm}.$$

2. A skála függvények ortogonálisak a waveletekre nézve:
$$\int_{-\infty}^{\infty} \varphi(x-n)\psi(x-m) = 0.$$

3. A $\psi_n^k(x) = 2^{n/2}\psi(2^n x - k)$ alakú waveletek minden skálázási szinten ortonormáltak egymásra nézve:
$$\int_{-\infty}^{\infty} \psi_n^k(x)\psi_m^\ell(x) = \delta_{nm}\delta_{k\ell}.$$

Tehát megmutattuk, hogy a multirezolúció elmélete miként hozható kapcsolatba a sávszűrést végző szűrőkkel. A továbbiakban bemutatjuk a wavelet-transzformáció gyorsítását, az úgynevezett *Gyors Wavelet-Transzformációt (Fast Wavelet Transform (FWT))* [2].

A wavelet-transzformációnak – csak úgy, mint a Fourier-transzformációnak

– az ortogonalitáson kívűl van még egy speviális tulajdonsága. Nevezetesen – mint láttuk – az n és $n-1$ skála szintek szoros kapcsolatban állnak egymással (az ω és $\omega/2$ frekvenciákkal analóg módon). Ily módon a szűrő készletek egy fáját lehet létrehozni. A D felülvágó szűrő a bemenet különbségét számolja, majd a $(\downarrow 2)$ downsampling művelete megtartja a páros sorszámú $x(2k)-x(2k-1)/\sqrt{2}$ különbségeket. Ezek végleges eredmények, a kapott értékeket nem szűrjük tovább. Az ezekből meghatározott b_{nk} wavelet-együtthatók a logaritmikus fa levelét reprezentálják. A C alulvágó szűrő lokális átlagokat számol, majd a $(\downarrow 2)$ downsampling művelete innen is kiszűri a páratlan sorszámú sorokat. Azonban az így kapott $x(2k)+x(2k-1)/\sqrt{2}$ átlagok még nem feltélenül végleges értékek, ugyanis a C, illetve D szűrőkkel tovább szűrhetjük őket, megkapva ezzel a következő szinthez tartozó $a_{n-1,k}$ és $b_{n-1,k}$ skála függvény, illetve wavelet együtthatókat. A módszer a következő rekurzióra épül:

$$\begin{array}{ll} \textit{Átlagok (aluláteresztő szűrő)} & a_{n-1,k} = \frac{1}{\sqrt{2}}(a_{n,2k}) + a_{n,2k+1}, \\ \textit{Különbségek (felüláteresztő szűrő)} & b_{n-1,k} = \frac{1}{\sqrt{2}}(a_{n,2k}) - a_{n,2k+1}. \end{array} \tag{6.36}$$

Minden újabb szűrés egy finobb n szintről egy durvább $n-1$ szintre visz minket.

A gyors wavelet-transzformáció nem más tehát, mint a szűrők egy fája. A FWT mátrix reprezentációja során az A analízis mátrix, egyszerű átlag-különbség szűrő-mátrixok véges szorzatából áll:

$$A = \begin{bmatrix} L & \\ B & \\ & I \end{bmatrix} \begin{bmatrix} L \\ B \end{bmatrix}. \tag{6.37}$$

A fenti formula jobb oldalán álló mátrix az első a fában, míg az eggyel balra lévő a rekövetkező szint alul-, illetve felüláteresztő szűrőit tartalmazza. A jobb alsó sarkában egy I egység-mátrix van, azzal a céllal, hogy a különbségek már ne legyenek újraszűrve. Az is észrevehető, hogy a mátrix nagyon sok nullát tartalmaz, amelyek a műveletvégzések számát jelentősek csökkentik. Természetesen – amennyiben az eredeti mátrixunk $l = 2^N$ hosszú volt – ezt a felbontást tovább lehetni folytani, összesen N lépésen keresztül. Ebben az esetben N szintje lenne a fának. A jobb oldali mátrix minden sora annyi nem nulla elemet tartalmaz, ahány együtthatója van az adott szűrőnek, jelöljük ezt a számot T-vel. Tehát a legfinomabb szinten a mátrixban szereplő nem nulla elemek száma Tl. A következő szint mátrixa már csak $Tl/2$ nem nulla elemet tartalmaz, ugyanis a feldolgozandó bemenet is feleakkora. Tehát N szinten az összes nem nulla elemek száma:

$$Tl\left(1 + \frac{1}{2} + \frac{1}{4} + \cdots + + \frac{1}{2^{N-1}}\right) < 2Tl,$$

melynek következményeképpen a gyors wavelet-transzformáció a $b = Ax$ N darab együtthatóját kevesebb, mint $2Tl$ szorzással határozza meg.

Hasonló számú művelet szükséges az $x = Sb$ szintézis elvégzéséhez, azonban az S-et alkotó mátrixok szorzásának sorrendje fordított. Tehát a szintézisfa az inverz mátrixokból áll *fordított sorrendben.*

6.4. Alkalmazások

Ebben a szkaszban bemutatunk néhány, jelfeldolgozásban használatos alkalmazást. A módszerek a wavelet analízis elemeire és a waveletek speciális tulajdonságaira épülnek.

Az első alkalamazásunk, a *waveletek zsugorítása (wavelet shrinkage)* Johnstone és Donoho nevéhez fűződik. Egy jel L-szintű wavelet dekompozíciójában a szignifikáns energiájú wavelet-együtthatók száma csekély. Egy jelet kevés számú együtthatóval pontosan lehet ábrázolni. A zsugorítás alapja, hogy az eljárás során csak bizonyos *küszöbértéknél (threshold)* magasabb értékű együtthatókat tartunk meg. A wavelet zsugorító algoritmus L szintre bontja a jelet, majd minden szinten kiválaszt egy küszöbértéket és alkalmazza az úgynevezett *kemény küszöbölés (hard thersholding)* műveletét. A küszöbölési technika annyit jelent, hogy a küszöbérték alatti elemek értékét nullával helyettesítjük. Ez által az eredeti jelet visszaállítani pontosan ugyanabba a formába nem lehet. Azonban az együtthatók meglepően sok hányadát ki lehet így nullázni, anélkül, hogy a visszaállított jelben különösebb – szemmel látható – minőségromlás keletkezzen. A nullázások miatt a transzformáltat leíró mátrix jól tömöríthető egy veszteségmentes tömörítő eljárással, például Huffman kódolással. Egy másik megoldás, a δ szinten történő *lágy küszöbölés (soft thresholding).* Az alábbiakban megadtuk egy $x(t)$ jel kemény (6.38), illetve lágy (6.39) küszöbölésének kimenetét:

$$y_h(t) = \begin{cases} x(t), & |x(t)| > \delta, \\ 0, & |x(t)| \leq \delta \end{cases} \text{ kemény (hard) küszöbölés,} \tag{6.38}$$

illetve

$$y_s(t) = \begin{cases} sign(x(t))(|x(t)| - \delta), & |x(t)| > \delta, \\ 0, & |x(t)| \leq \delta \end{cases} \text{ lágy (soft) küszöbölés.} \tag{6.39}$$

A következő alkalmazásunk a *zajmentesítés (denoising)* folyamata. A küszöbölés általában az eredeti jel aluláteresztő változatát adja. A δ köszöbérték megfelelő megválasztásával lehetőség nyílik a $w(n)$ zaj kiszűrésére az eredeti $x(n) = u(n) + \sigma w(n)$ jelben. A legegyszerűbb esetben a $w(n)$

függvény az úgynevezett *Gauss-féle fehér zaj*, ami normális eloszlást követ a $(0, 1)$ intevallumon, más szóval $w(n) \approx N(0, 1)$. A példa kedvéért tegyük fel, hogy a jel és a zaj ismeretlen függvények, a bemenő adatok száma $N = 2^{J+1}$, és a zaj szintje σ ismeretlen. Ekkor a zajmentesítés a következő lépésekből tevődik össze:

1. lépés: Bontsuk fel a zajos $x(n)$ jeleket, megkapva ezzel a zajos ψ_n^k waveleteket ($n = n_0, \ldots, J$, $k = 0, \ldots, 2^n - 1$).

2. lépés: Válasszuk a köszöbértéket a következőképpen:

$$\delta_n = \frac{\sqrt{2\log(n)}\sigma}{\sqrt{(n)}}$$

és alkalmazzuk a lágy (soft) küszöbölési technikát $\delta = \delta_n$ küszöbértékkel.

3. lépés: Becsüljük meg a $\widehat{b}_{nk}$ wavelet-együtthatókat.

4. lépés: Állítsuk nullára az összes olyan $\widehat{b}_{nk}$ wavelet-együtthatót, ahol $n > J$ és hajtsuk végre az inverz wavelet-transzformációt,melynek eredménye a becsült $\widehat{u}(n)$ függvényt.

A következőekben az *éldetektálásról (edge detection)* lesz szó. Egy jelben az él nem más, mint az intenzitásban bekövetkező drasztikus változás (legyen szó akár egy-, akár kétdimenziós jelről), mintegy megtörve annak folytonosságát. Egy zajos wavelet-együtthatókból álló szorzat, különböző skála szinteken, alkalmas arra, hogy éleket határozzunk meg. Él detektálásához használjuk $n = 1$ és $n = 2$ skálákat, ekkor a C_{12} szorzat a downsampling művelete nélkül alkalmazott wavelet-transzformáció két együtthatója – b_{1k} és b_{2k} – között helyezkedik el. Ekkor C_{12} energiája az újraskálázás miatt megegyezik b_{1k} energiájával, nevezetesen

$$\sum_k (b_{1k})^2 = \sum_k (C_{12})^2 = \sum_k (b_{1k}b_{2k})^2.$$

A továbbiakban a független k eltolás indexét – az egyszerűség kedvéért – mellőzzük. Tehát egy él pontosan azon a helyen található, ahol $|C_{12}| > |b_1|$. Ezek a helyek egy speciális, bináris maszkba elmentjük, majd a C_{12}, illetve a b_1 értékek nullára állítjuk. Ezek után C_{12} energiáját újraskálázzuk, melynek következtében az új szorzat már C_{12} és b_1 között fog elhelyezkedni. Erre megint megvizsgáljuk az éleket, amelyek helyét megint elmentjük. Ezt az eljárást addig folytatjuk, amíg a b_1 energiája el nem ér egy bizonyos zajos szinet b_1-ben.

Most megmutatjuk, hogy a waveletek miként alkalmazhatóak a kétdimenziós képek feldolgozására. Minden kép felfogható tulajdonképpen egy A mátrixnak, ahol az a_{ij} elemek értékei, a kép (i,j) pozícióján lévő pixelértékek. Az egyszerűség kedvéért tegyük fel, hogy ezek az értékek a szürke-skálán (greyscale) adottak, azaz $a_{i,j} \in \{0,\ldots,255\}$. Továbbá tegyük fel azt is, hogy a kép mérete – tehát a mátrix dimenziója – $2^n \times 2^n$, ahol n teszőleges egész. Ekkor a wavelet-dekompozíciós eljárás a következőképpen alakul:

1. Az A mátrix soraira alkalmazzuk a H aluláteresztő és a G felüláteresztő szűrőket. Ennek következményeképp két mátrix keletkezett, jelöljük ezeket H_rA és G_rA szimbólumakkal, ahol a r index azt hivatott mutatni, hogy a szűrést a sorokon (row) végeztük el. A kapott mátrixok dimanziója $2^n \times 2^{n-1}$.

2. Most alkalmazzuk e két – H_rA és G_rA – mátrix oszlopaira (column) a H és G szűrőket. Eredményül négy $2^{n-1} \times 2^{n-1}$ dimenziójú mátrixot kapunk: H_cH_rA, H_cG_rA, G_cH_rA és G_cG_rA. ezek közül az első H_cH_rA mátrix tartalmazza az átlagot, míg a többi három a részleteket.

3. A feldolgozás folytatható – a gyors Wavelet-transzformációnál látottak alapján – a következő szint szűrésével. Azonban most a kiinduló mátrixunk az átlagolt H_cH_rA mátrix (hiszen az FWT során is az átlagokat szűrtük tovább).

E kétdimenziós dekompozíciós eljárásnak – és persze a jól lokalizálhatóságnak – köszönhetőek a wavelet-transzformáció legnagyobb sikerei. Ugyanis ezen eljárás segítségével könnyen lehet hatalmas méretű képeket *tömöríteni*. A sikerességet nem kicsinyíti az a tény, hogy módszer a JPEG2000 elfogadott szabványa lett [2].

Az éldetektáló algoritmus is kiterjeszthető kétdimenziós képekre. Először elvégezzük a wavelet-transzformációt mind horizontálisan, mind vertikálisan, majd a horizontális és vertikális éleket külön-külön cimkézzük W_{cv} és W_{ch} szimbólumokkal. Az élek bínáris tárolását is két – W_h és W_v – maszkban végezzük, amelyeket a W_{ch} és W_{cv} értékekből határozunk meg küszöböléssel (küszöbnek vmilyen zajos értéket választva). Az élek horizontális M_h (illetve vertikális M_v) irányait tartalmazó képet, a bínáris W_h (illetve W_v) élmátrixok és a legalacsonyabb szinten lévő b_{1h} (illetve b_{1v}) wavelet-együtthatók szorzásával kaphatjuk meg. Majd végül e két mátrixból már megkaphatjuk a teljes képre vonatkozó együttes él információkat tartalmazó mátrixot:

$$M(x,y) = \sqrt{|M_h|^2 + |M_v|^2}.$$

6.5. Fourier-Transzformáció kontra Waveletek

Ebben a szakaszban összehasonlítjuk a jelfeldolgozás két igen fontos – és nélkülözhetetlen – transzformációját, rávilágítva az egyes hasonlóságokra, illetve különbségekre. Ezek a transzfomációk természetesen a Fourier- (FT), illetve Wavelet-transzformációk (WT). Kezdjük az elemzést a hasonlóságokkal.

A legszembetűnőbb kapcsolat a két eljárás között, hogy mind a FT, mind a WT egy líneáris operátor, ami invertálható és ortogonálissá tehető. Mindkét transzformáció tulajdonképpen egy adott függvénytér (idő-tartomány) elforgatását végzi egy másik tartományra. Legyen szó bármelyik transzformáció teréről, véve onnan egy függvényt, az felírható az kapott új tér bázis függvényeivel vett líneáris kombinációs alakban. Míg a FT bázisai szinusz és coszinusz hullámok, addig a WT bázisai a waveletek.

Továbbá, áttérve a frekvencia-tartományra, észrevehető, hogy abban a térben mindkét transzfomáció *jól lokalizál*, melynek köszönhetően kialakultak olyan matematikai módszerek, mellyekkel például bizonyos – fölösleges – frekvencia-komponensek eltávolíthatók.

A transzformációk diszkrét változatai (FFT és DWT) a bemenetből egy $\log_{2^n}$ szegmensből álló adatszerkezetet generálnak és azt transzformálják át egy másik 2^n hosszúságú adat tömbbe. A mátrixok matematikai tölajdonságainak köszönhető a két transzformáció közötti igen nagy fokú hasonlóság. Ugyanis, például az FFT és DWT inverz-transzformációs mátrixai éppen az eredeti transzformációs mátrix transzponáltjaival egyeznek meg.

A legfontosabb különbség a két transzformáció között, hogy a *wavelet függvények lokalizálhatók időben is.* A Fourier-transzformációhoz tartozó szinusz és koszinusz függvényekre ez nem igaz, hiszen mindkét függvény az egész $(-\infty, \infty)$ intervallumon értelmezve van. Továbbá az FT bázis függvényei periodikusak, ez által lehetetlenné válik annak meghatározása, hogy az adott transzformáció melyik időpillanatban lett elvégezve. Ezzel szemben a wavelet-transzformációból ez az infomáció is kinyerhető. Ez az „ellenőrizhetőség" az alapja számos alkalmazásnak, mint például az adat tömörítési eljárások, amik képek speciális tulajdonságainak detektálását végzik, amelyek meghatározásával a zajos elemek kiszűrhetőek az időtartományban.

A FT időben való lokalizációjára a megoldást *Gábor Dénes* vezette be. Az űgynevezett *Rövid-Idejű Fourier-Transzformáció (Short-Time Fourier Transzform STFT)* [7] úgy próbálja kiküszöbölni ezt a hiányosságot, hogy egyszerre csak az adott jel egy kis részletét transzformálja. Az STFT elvégéséhez szükség van egy $w(x)$ úgynevezett *ablak függvényt.* Ezen függvényeknek a tartója a $[-L, L]$ intervallum, azaz csupán $2L + 1$ hosszúságú szakaszon nem nulla az értékük. A teljes transzformációt ezen ablak függvény, időskálán való tologatásával lehet elvégezni. Az a időponthoz tartozó – gyakran *ablakozott Fourier-transzformációként* is emlegetett – STFT definíciója a következő:

$$X_{WT}(\omega, a) = \int_{-\infty}^{\infty} x(t)w(t-a)e^{-i\omega t}\, dt, \tag{6.40}$$

ahol $x(t)$ a bemenő jel. A legnagyobb előnye ennek a módszernek, hogy ha a bemenő jel legnagyobb energiája a $[-L, L]$ idő-intervallumon és $[-\Omega, \Omega]$ frekvencia-tartományon van, akkor STFT lokalizálni fog a $[-L, L] \times [-\Omega, \Omega]$ tartományon, és nullához közeli értékek lesznek a frekvencia- és időtartományon ott, ahol a jel kisebb aktivitást fejt ki. Azonban e módszer hátránya sem elhanyagolható, ugyanis a transzformáció során az ablak mérete nem változtatható. Ennek következtében minden frekvencia-komponensre ugyanazt az ablakot alkalmazzuk, tehát az analízis felbontása ugyanaz lesz minden helyen az idő-frekvencia síkon.

A WT előnye tehát az is, hogy nem rögzített méretű „ablakokkal" dolgozik. Az ablakokat ebben a transzformációban az anyawavelet és annak átskálázott és eltolt függvényei jelentik. Tehát a wavelet-transzformációnak végtelen sok lehetséges bázisfüggvényét lehet előállítani a már megismert módszerekkel. Ebben a felfogásban a waveleteket – a idő-frekvencia sík helyett – sokkal inkább az *idő-skála* síkon szokták emlegetni.

7. fejezet

Összefoglalás

A dolgozat a Fourier amalízis matematikai alapjainak ismertetésén túl bemutatta, milyen módon lehet a módszert a matematika más ágain és a gyakorlatiasabb jelfeldolgozás területén felhasználni. Kezdve egy általánosabb struktúrán, fokozatosan haladtunk a számunkra fontosabb speciális esetek felé, bemutatva ezzel az alkalmazni kívánt módszer jelentőségét, sokszínűségét. A transzformációt megvalósító gyors FFT algoritmusok bevezetésével és elemzésével rávilágítottunk az elmélet gyakorlati élteben való nélkülözhetetlen szerepére. Megmutattuk, hogy az eljárást miként lehet alkalmazni a jelfeldolgozás egyik klasszikus transzformációjára, a szűrésre, lecsökkentve ez által a műveletek elvégzéséhez szükséges futási időket. A módszer általánosításának tekintett z-transzformáció bevezetésével – a komplex függvénytan ismert eredményeinek segítségével – lefektettük a digitális jelfeldolgozás alapjait.

A Fourier analízis széles eszköztárának bemutatása után rátértünk egy viszonylag új témának számító, hasonló módszerekre épülő, nem kevésbé hatalmas jelentőségű tudományág ismertetésére, a waveletekre. A pontos matematikai definíciókon és levezetéseken keresztül, a fontosabb összefüggések bemutatására segítségül hívtuk, a már megismert Fourier elméletbeli módszereket, szemléltetve ezzel, annak – wavelet analízisbeli – nélkülözhetetlenségét. Ezen új elmélet alapjául szolgáló multirezolúciós analízis fontosabb elemeinek ismertetése után, a skálázási egyenlet levezetésével, megoldásával bevezettük a jelfeldolgozás legfontosabb elemeit, a szűrőket.

A dolgozat, tartalmilag utolsó fejezetében betekintést nyerhettünk a digitális jelfeldolgozás alapjaiba. Ennek során a már megismert módszereket alkalmaztuk, e gyakorlatban egyre szélesebb köben eltejedő tudományág bemutatására, felépítésére. A már – matematikailag – megismert aluláteresztő szűrő gyakorlati oldaláról való megközelítése mellett bevezettük annak felüláteresztő párját. A klasszikus Haar példán keresztül szemléltettük azok hatását – a már megismert módszerek segítségével – a különböző – idő és frekvencia – tartományokban. Bemutattuk azt is, hogy miként lehet e két szűrőt egy 2-csatornás szűrőkészletben egyszerre alkalmazni, úgy, hogy a bemenetként megadott jel tökéletesen rekonstruálható legyen. A downsampling és upsampling műveletek bevezetésével, a szűrés során keletkező jelek – nem elhanyagolható – tárolási igényeinek csökkentését tűztük ki célul. A wavelet analízisbeli multirezolúciót kapcsolatba hoztuk szűrőinkkel, melynek köszönhetően képesek lettünk egy – szűrést megvalósító – igen gyors algoritmus bevezetésére. Az elmélet, gyakorlati oldaláról való megközelítése céljából bemutattunk néhány, jelfeldolgozásban kiemelkedő jelentőségűnek számító, alkalamzást, mint például a zajmentesítést, vagy a tömörítést.

A Fourier analízis kiemelkedő jelentőségének és más tudományágban való nélkülözhetetlen szerepének ismertetése mellett fontosnak tartottuk megemlíteni annak hiányosságát, hátrányát. A wavelet-transzformációval való összehasonlító elemzés során kiderülnek a módszer hézagai, gondoljunk csak például az időben való lokalizálhatóság kérdésére. Azonban, ezen apró hibák figyelembe vételével is kijelenthetjük, hogy a módszer korszakalkotónak számít. A Fourier analízis elmélete számos tudományág rohamos fejlődésének adta meg a kezdőlökést, kivívva ezzel magának az elismerést.

Irodalomjegyzék

[1] Dr. Simonyi Ernő: *Digitális szűrők*, Műszaki Könyvkiadó, 1984, [563], ISBN 963 10 5306 7.

[2] G. Strang and T. Nguyen: *Waveletes and Filter Banks*, Wellesley-Cambridge Press, 1996, [512], ISBN 09614088-71.

[3] Schipp Ferenc: *Fourier-Analízis*, ELTE Egyetemi jegyzet, 2006, [121].

[4] F. Scipp and W. H. Wade: *Transforms on Normed Fields*, Janus Pannonius University, Pecs, 1995, [176].

[5] Fehér Tamás: *Wavelet-Transzformáció a Képfeldolgozásban*, Informatika a Felsőoktatásban, 2005, [6].

[6] Brani Vidakovic and Peter Mueller: *Wavelets for Kids*, Duke University, [28].

[7] Martin Vetterli: *Waveletes and Filter Banks: Theory and Design*, IEEE Transactions on Signal Processing, Vol 40, 1992, [26].

[8] Schipp Ferenc: *Waveletek*, ELTE Egyetemi jegyzet, 2003, [90].

[9] T. H. Cormen C. E. Leiserson R. L. Rivest C. Stein: *Új algoritmusok*, Scolar Kiadó, 2003, [992], ISBN 963 9193 90 9.

[10] Pál J. Schipp F. Simon P.: *Analízis II.*, Nemzeti Tankönyvkiadó, 1997, [267].

[11] Járai Antal: *Bevezetés a Matematikába*, ELTE Eötvös Kiadó, 2007, [254], ISBN 963 463 729 9.

[12] Hanka L. Zalay M.: *Komplex függvénytan*, Műszaki Könyvkiadó, 2003, [416], ISBN 963 16 2816 7

[13] *http://itl7.elte.hu/html/jelfel/jelfeld.htm*, 2007.03.16.

[14] *http://www.amara.com/IEEEwave/IEEEwavelet.html*, 2007.04.04.

[15] *http://perso.orange.fr/polyvalens/clemens/wavelets/wavelets.html*, 2007.04.04.

[16] *http://www.engmath.dal.ca/courses/engm6610/notes/notes.html*, 2007.05.01.

[17] *http://mazsola.iit.uni-miskolc.hu/DATA/segedletek/digszurok.doc*, 2007.03.16.

[18] *http://www.mathworks.com/access/helpdesk/help/toolbox/wavelet/*, 2007.04.04.

[19] *hu.wikipedia.org/wiki/Digitális-jelfeldolgozás*, 2007.03.20.

[20] *http://en.wikipedia.org/wiki/Wavelet*, 2007.03.20.