COMP.SGN.100 Signaalinkäsittelyn perusteet

Sari Peltonen

Signaalinkäsittely
Tietotekniikan yksikkö
Tampereen yliopisto

Digitaalinen signaalinkäsittely

- Digitaalisesta signaalinkäsittelystä on tullut yksi nykytekniikan avainaloista, ja se tukee läheisesti ainakin tietoliikennetekniikkaa, mittaustekniikkaa ja tietotekniikkaa.
- Digitaalisen signaalinkäsittelyn (Digital signal processing, DSP) voidaan katsoa syntyneen 1960–1970–luvuilla, jolloin tietokoneet alkoivat olla riittävän yleisesti käytettävissä.
- Tämän jälkeen sitä on menestyksekkäästi sovellettu lukuisilla alueilla; lääketieteellisestä PET-kuvantamisesta CD-soittimeen ja GSM-puhelimeen.
- Sovellukset ovat varsin lukuisat, joten kaikkea DSP:stä on mahdotonta hallita (eikä niitä ole mielekästä opettaa korkeakoulussa).
- Tärkeimmät perusmenetelmät ovat kuitenkin pysyneet vuosien varrella samoina.

Digitaalinen signaalinkäsittely

- Jatkossa on tarkoitus käsitellä
  - tärkeimmät peruskäsitteet,
  - osa tärkeimmistä menetelmistä,
  - esimerkkisovelluksia.
- Kun lineaaristen järjestelmien perusasiat on käsitelty, perehdytään tyypilliseen signaalinkäsittelyn ongelmaan: kuinka poistaa tietyt taajuudet annetusta signaalista.
- Tulevilla kursseilla perehdytään tarkemmin signaalinkäsittelyn menetelmiin sekä tarkastellaan sovelluskohteita lähemmin.

Mitä signaalinkäsittelyllä tarkoitetaan

- Tyypillinen DSP-sovellus sisältää seuraavat vaiheet:
  1. Niin sanottu A/D-muunnin (analog/digital) muuntaa vastaanotetun (jatkuva-aikaisen) analogisen signaalin digitaaliseksi ja diskreettiaikaiseksi.
  2. Tämän jälkeen diskreettiaikaista digitaalista signaalia muokataan jollain järjestelmällä (esim. tietokoneella). Tätä vaihetta kutsutaan **suodattamiseksi**. Suodatuksen tavoite on muuntaa järjestelmään saapuva signaali sovellutuksen kannalta hyödyllisempään muotoon. Tämä saattaa tarkoittaa esimerkiksi:
     - Signaalissa olevan kohinan poistamista siten, että varsinainen signaali säilyy mahdollisimman hyvin.
     - Signaalissa olevien mielenkiintoisten piirteiden erottelua muun signaalin joukosta.

Mitä signaalinkäsittelyllä tarkoitetaan

Perinteisesti suotimet ovat olleet lineaarisia niiden helpomman toteuttamisen ja analysoinnin vuoksi, mutta myös epälineaarisia suotimia on tutkittu.

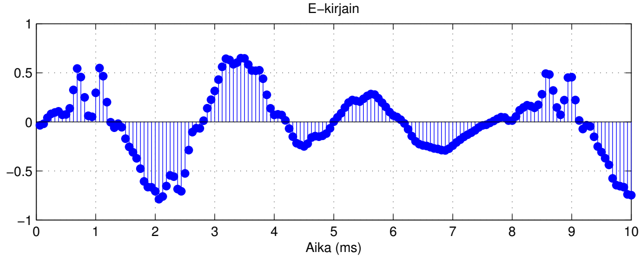
$$x(n) \longrightarrow \boxed{\text{Digitaalinen suodatin}} \longrightarrow y(n)$$

3. Suodatuksen jälkeen signaali muunnetaan takaisin analogiseksi D/A-muuntimella.

- Jatkossa keskitytään pääasiassa vaiheeseen 2, suotimen suunnitteluun.

Mitä signaalinkäsittelyllä tarkoitetaan

- Käsiteltävä signaali voi esittää esimerkiksi ääntä, puhetta, pulssia, aivokäyrää, maanjäristystä, pörssikursseja tai mitä hyvänsä mitattavissa olevaa aikasarjaa.
- Alla olevassa kuvassa on 10 millisekunnin mittainen näyte puhesignaalista:

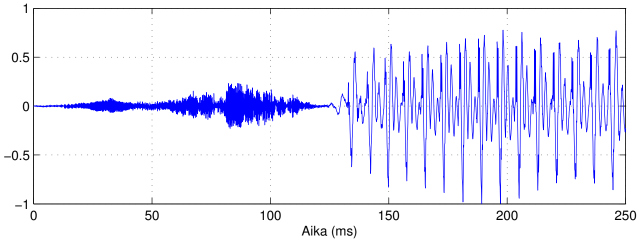


Mitä signaalinkäsittelyllä tarkoitetaan

- Mikrofoni muuntaa mittaamansa pienet ilmanpaineen vaihtelut sähköiseen muotoon, josta tietokone muuntaa ne edelleen digitaaliseen muotoon tallentamalla jännitteen hetkelliset lukuarvot $1/16000$ sekunnin välein.
- Tässä tapauksessa **näytteenottotaajuus** on siis 16000 hertsiä.

Mitä signaalinkäsittelyllä tarkoitetaan

- Seuraavassa kuvassa on pidempi näyte samasta signaalista:



- Kuvan signaali sisältää sanan ”seitsemän” kaksi ensimmäistä kirjainta.
- S-kirjain sijaitsee alusta lukien noin 125 millisekunnin matkalla, ja seuraavassa 125 millisekunnissa on E-kirjain.

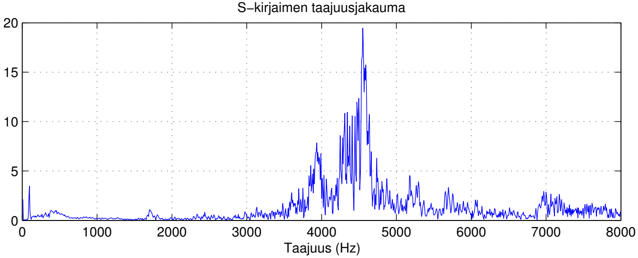
Mitä signaalinkäsittelyllä tarkoitetaan

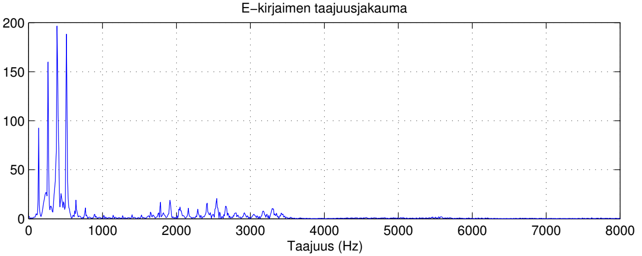
- Konsonantin ja vokaalin ero näkyy hyvin: soinnillisen vokaalin kohdalla on selkeä ylös-alas-värähtelykuvio ja konsonantin kohdalla lukuarvot ovat satunnaisempia.
- Useille erityyppisille signaaleille on luontevaa ajatella niiden koostuvan yksittäisistä taajuuksista (yksittäisistä sinisignaaleista sopivassa suhteessa).
- Esimerkiksi äänisignaaleita on helpoin ymmärtää ja analysoida niiden taajuusjakauman kautta.
- Kuten myöhemmin kurssilla tulemme näkemään, taajuusjakauma voidaan laskea **Fourier-muunnoksen** avulla.

Tampereen yliopisto

Mitä signaalinkäsittelyllä tarkoitetaan

- Alla olevissa kuvissa on laskettu edellisen kuvan S- ja E-kirjainten taajuusjakaumat:





Mitä signaalinkäsittelyllä tarkoitetaan

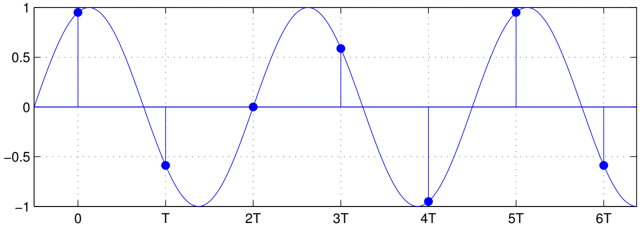
- Kuvista näkyy selvästi, että
  - S-kirjaimen sisältämät taajuudet jakautuvat melko laajalle alueelle sekä melko suurillekin taajuuksille,
  - kun E-kirjain sisältää vain yksittäisiä pieniä taajuuksia (korkeat piikit).

Näytteenottoteoreema

- Edellisessä esimerkissä näytteenottotaajuus oli 16000 hertsiä eli 16 kHz.
- Mitä suurempi näytteenottotaajuus on, sitä pienemmät yksityiskohdat signaalista saadaan talteen.
- Suurempi näytteenottotaajuus vaatii kuitenkin enemmän tilaa, joten taajuutta ei kannata nostaa liian suureksi.
- Mistä siis tiedetään mikä on riittävä näytteenottotaajuus tietylle signaalille?
- Tähän kysymykseen vastaa **näytteenottoteoreema, engl. sampling theorem.**

Näytteenottoteoreema

- Jatkuva-aikaista signaalia näytteistettäessä siitä otetaan näytteitä ajanhetkillä 0, T, $2T$, $3T, \ldots$ ja vain signaalin näillä hetkillä saamat arvot talletetaan (siniset ympyrät alla olevassa kuvassa):

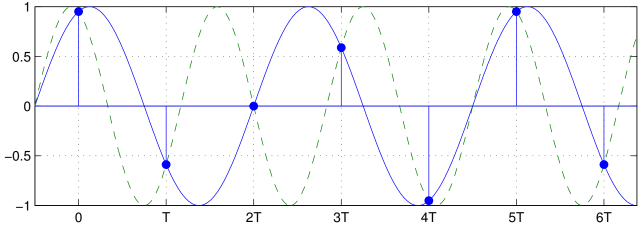


Näytteenottoteoreema

- Jos siis jatkuvasta signaalista käytetään merkintää $x_c(t)$, missä $t \in \mathbf{R}$, niin näytteistyksen tuloksena saadaan lukujono $x(n)$, jolle on voimassa ehto $x(n) = x_c(nT)$ ($n = 0, 1, 2, \ldots$).
- Vakio T ilmoittaa siis kuinka monta sekuntia on kahden peräkkäisen näytteen väli.
- Useimmiten sama asia ilmaistaan sanomalla montako kertaa sekunnissa näytteitä otetaan.
- Tämän suureen nimi on **näytteenottotaajuus** (sampling frequency tai sampling rate) ja se on vakion T käänteisluku, $F_s = \frac{1}{T}$.
- Jos näytteenottotaajuus on liian pieni (ja siis näytteiden väli T liian suuri), tapahtuu **laskostumista** eli **alinäytteistymistä** (aliasing).

Näytteenottoteoreema

- Laskostuminen tulee ilmi alla olevasta kuvasta:



- Kuvan kahdella sinisignaalilla on samat näytearvot, koska näytteenottotaajuus on liian pieni suurempitaajuuksiselle signaalille (katkoviiva).
- Muunnettaessa näin näytteistettyä signaalia takaisin analogiseksi lopputuloksena on pienempitaajuinen sinisignaali (yhtenäinen viiva).
- Sanotaan, että suurempi taajuus laskostuu pienemmän taajuuden päälle.

Näytteenottoteoreema

- Pohdittaessa riittävää näytetaajuutta edellisen esimerkin sinisignaalille tuntuu, että kaksi näytettä jaksoa kohti saattaisi riittää.
- Tällöin nimittäin tallennettaisiin signaalin suurin ja pienin arvo ja näiden avulla voitaisiin interpoloida muut lukuarvot huippuarvojen välille.
- Luonnollisesti mikä tahansa tätä suurempi taajuus käy yhtä hyvin.
- **Näytteenottoteoreema** kertoo, että tämä arvaus pitää paikkansa.:

> *Jatkuva-aikainen signaali voidaan muodostaa uudelleen näytearvoistaan, jos näytteenottotaajuus F_s on vähintään kaksi kertaa niin suuri kuin signaalin sisältämä suurin taajuuskomponentti.*

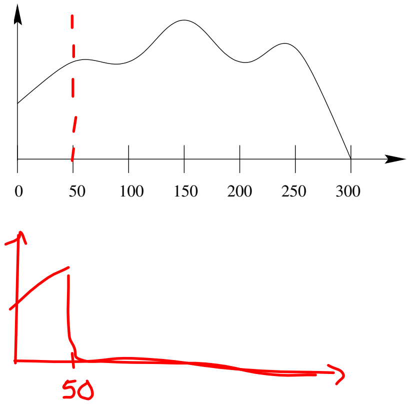
Näytteenottoteoreema

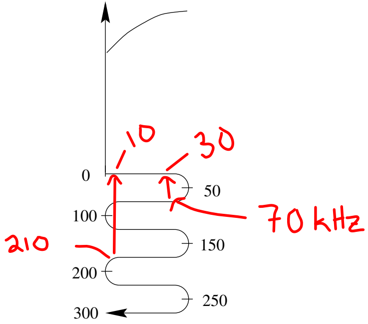
- Jos edellinen ehto ei ole voimassa, täytyy taajuutta $F_s/2$ suuremmat taajuudet leikata pois jollain analogisella järjestelmällä laskostumisen estämiseksi.
- Taajuudesta $F_s/2$ käytetään nimitystä Nyquistin taajuus (Nyquist frequency, Nyquist rate).
- Näin ollen enintään 300 kHz taajuuksia sisältävä signaali vaatii järjestelmän, jonka näytteenottotaajuus on vähintään 600 kHz.
- Jos näin ei ole käy kuten seuraavassa kuvassa.
- Kuvat esittävät eräiden signaalien spektrejä, jotka kertovat, kuinka paljon kutakin taajuutta on mukana signaalissa.

Tampereen yliopisto

Näytteenottoteoreema

- Kuvan alkuperäisessä jatkuvassa signaalissa on taajuuksia 300 kHz asti (vasemmalla), jolloin 100 kHz taajuudella näytteistettäessä yli 50 kHz taajuudet summautuvat alemmille taajuuksille laskostuneen x-akselin mukaisesti (oikealla).





Näytteenottoteoreema

- Lopputuloksessa esimerkiksi taajuudella 25 kHz on usean taajuuskomponentin summa: 25 kHz, 75 kHz, 125 kHz, 175 kHz, 225 kHz ja 275 kHz.
- Toisaalta 10 kHz taajuuden kohdalla ovat seuraavat taajuudet summautuneena: 10 kHz, 90 kHz, 110 kHz, 190 kHz, 210 kHz ja 290 kHz.
- Kyseisten kuuden komponentin summasta ei voida palauttaa enää alkuperäistä signaalia.
- Paras tulos esimerkin 100 kilohertsin näytteenottotaajuudella saadaan poistamalla yli 50 kilohertsin taajuudet ennen näytteenottoa.
- Vaikka taajuudet 50 kHz – 300 kHz menetetäänkin, saadaan edes taajuudet 0 kHz – 50 kHz tallennettua alkuperäisessä muodossaan.

Digitaalisen signaalinkäsittelyn etuja ja haittoja jatkuva-aikaisiin suodattimiin nähden

- Analogisia suotimia on elektroniikan alalla tutkittu jo kauan.
- Nämä kootaan elektronisista komponenteista ja ne poimivat tyypillisesti tietyt taajuudet signaalista ja poistavat muut.
- Herää kysymys, miksi sama pitäisi tehdä digitaalisesti?
- Digitaalisista suotimista on helppo tehdä tarkkoja, ja niiden ominaisuudet pysyvät samoina koko käyttöajan.
- Digitaalisten suodinten ominaisuudet määräytyvät niiden kertoimien kautta, eivätkä tietokoneohjelman kertoimet muutu esimerkiksi ajan myötä tai lämpötilan vaihdellessa.
- Tarkkuus saadaan myös helposti paremmaksi lisäämällä laskentatehoa ja laskentatarkkuutta.

Digitaalisen signaalinkäsittelyn etuja ja haittoja jatkuva-aikaisiin suodattimiin nähden

- Analogisistakin järjestelmistä voidaan toki tehdä yhtä tarkkoja ja ominaisuutensa säilyttäviä, mutta tällöin on käytettävä kalliimpia ja laadukkaampia komponentteja.
- Usein sanotaankin, että digitaalinen CD-soitin toi HIFI-laadun tavallisen kuluttajan ulottuville, kun analogisilla laitteistoilla se oli ainoastaan varakkaiden saatavilla.
- Digitaalisilla suotimilla on useita teorian kannalta hyviä ominaisuuksia.
- Niiden avulla voidaan esimerkiksi toteuttaa täysin **lineaarivaiheinen** suodin, mikä on mahdotonta analogisen suotimen avulla.
- Lineaarivaiheisuus tarkoittaa sitä, että kaikki signaalin sisältämät taajuudet viivästyvät yhtä paljon.

Digitaalisen signaalinkäsittelyn etuja ja haittoja jatkuva-aikaisiin suodattimiin nähden

- Lisäksi digitaaliset suotimet toimivat kuin tietokoneohjelmat, joten niihin voidaan lisätä monimutkaisiakin rakenteita, joita analogisilla järjestelmillä on mahdoton toteuttaa.
- Tärkein syy digitaalisten suodinten käyttöön analogisten komponenttien sijaan on kuitenkin raha: samaa signaaliprosessoria voidaan käyttää useisiin eri sovelluksiin, jolloin sitä voidaan tuottaa suuremmissa erissä ja suuret tuotantoerät painavat prosessorien hintoja alas.
- Prosessoreja käyttävät yritykset puolestaan toteuttavat oman tuotteensa ohjelmistona fyysisten laitteiden sijaan.
- Tällöin tuotteen monistaminen on helppoa, ja sama tuote voidaan myydä useaan kertaan—aivan kuten tietokoneohjelmistotkin.

Digitaalisen signaalinkäsittelyn etuja ja haittoja jatkuva-aikaisiin suodattimiin nähden

- Toisaalta hyvin yksinkertaiset järjestelmät, jotka eivät tarvitse suurta tarkkuutta, on helpointa toteuttaa analogisilla komponenteilla.
- Digitaalinen järjestelmä tarvitsee aina A/D ja D/A-muuntimet sekä prosessorin.
- Jos tavoitteena on vain jakaa autostereoiden kaiutinsignaali kahteen eri taajuuskaistaan, ei tätä varten kannata rakentaa digitaalista järjestelmää.

Sovelluskohteita

Seuraavaksi tarkastellaan muutamia tyypillisiä sovelluskohteita melko pintapuolisesti.

Kompressio:

- Kompressioalgoritmit pyrkivät poistamaan näytteiden välillä olevaa **redundanssia**, ja saamaan näin tarvittavien bittien määrää vähennetyksi.
- Erilaisia pakkausmenetelmiä on lukuisia.
- Menetelmä saattaa olla suunniteltu pakkaamaan tehokkaasti esimerkiksi äänisignaalia, digitaalista kuvaa tai videota tai vaikkapa lääketieteessä tavattavia kolmiulotteisia tomografiakuvia.
- Menetelmät käyttävät hyväkseen kunkin sovelluksen ominaispiirteitä.

Sovelluskohteita

Kompression kohokohta:

- Keväällä 2017 JPEG-komitea (joka valvoo kuvankompressiostandardien kehittämistä) käynnisti kansainvälisen hankkeen plenoptisten kuvien kompressiostandardin kehittämiseksi.
- Professori Ioan Tabusin tutkimusryhmän (Signaalinkäsittely, Tampereen yliopisto) ehdotus valittiin parhaaksi ja standardia rakennetaan sen pohjalta.

Sovelluskohteita

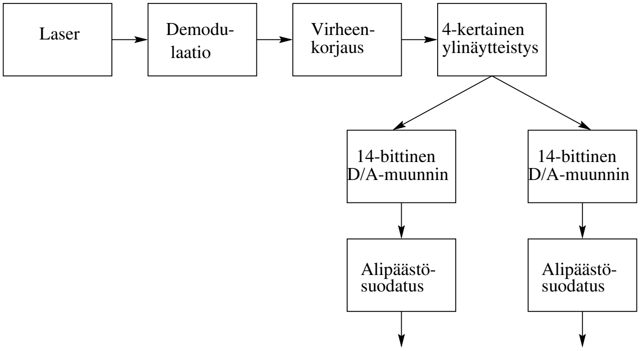
Kaiunkumous:

- Koska puhelinliikenteen signaalit kulkevat kahteen suuntaan, ongelmaksi saattaa muodostua käyttäjälle ärsyttävä kaiku.
- Jokainen puhuttu sana palaa pienellä viiveellä linjaa pitkin takaisin ja saa aikaan kaikua vastaavan efektin.
- Erityisenä ongelmana tämä on kaiutinpuhelinta käytettäessä, mutta myös tavallisissa puhelimissa.
- Ongelma voidaan poistaa riittävän tehokkaasti adaptiivisen signaalinkäsittelyn menetelmin.

Sovelluskohteita

CD-soitin:

- Alla oleva yksinkertaistettu kaavio tyypillisestä CD-soittimesta sisältää lukuisia kohteita, joissa käytetään DSP:tä.



Sovelluskohteita

Puheentunnistus:

- 2000-luvun alussa puheentunnistusta tehtiin perinteisillä lähestymistavoilla, kuten esim. Hidden Markov -malli yhdistettynä eteenpäin syöttäviin (feedforward) neuroverkkoihin.
- Viimeaikaisia syväoppimisen ja big datan edistysaskeleita on hyödynnetty myös puheentunnistuksessa, erityisesti käyttäen syväoppimismenetelmää, jossa takaisinkytketyn neuroverkon (recurrent neural network, RNN) arkkitehtuurina on pitkä lyhytkestomuisti (Long Short-term memory, LSTM).
- Vuonna 2015 Googlen puheentunnistukseen saatiin dramaattinen 49 % suorituskyvyn nousu CTC (Connectionist Temporal Classification) -koulutetun LSTM:n avulla.

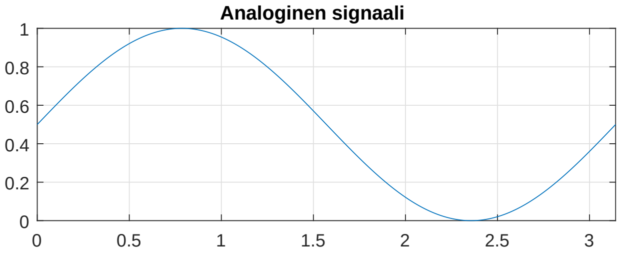
Sovelluskohteita

Tietokonetomografia:

- Lääketieteellisessä kuvantamisessa käytetään tekniikkaa funktionaalisten tai anatomisten kuvien ottamiseksi elävän kehon sisäpuolelta.
- Lääketieteellisen kuvantamisen tavoitteena on antaa kuva kehon sisäpuolelta tavalla, joka on mahdollisimman ei-invasiivinen.
- Esim. tietokonetomografia, MRI, PET, SPECT ja elektronitomografia ovat tomografisen kuvantamisen modaliteetteja.
- Tomografisessa kuvantamisessa kuva rekonstruoidaan projektiodatasta, joka usein esitetään sinogrammina.

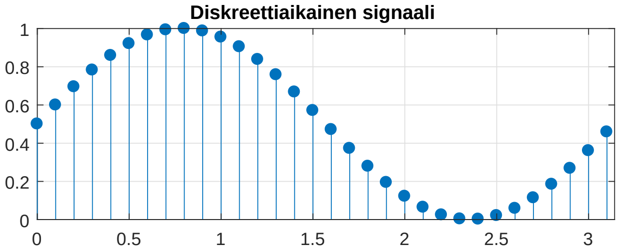
Signaalit ja järjestelmät

- Edellä näimme jo esimerkkejä digitaalisista signaaleista.
- Määritellään nyt käytettävät käsitteet täsmällisesti.
- **Analoginen signaali** on määritelty jokaisella ajanhetkellä ja se voi saada äärettömän määrän eri arvoja (esim. väliltä $[0, 1]$, kuten alla).



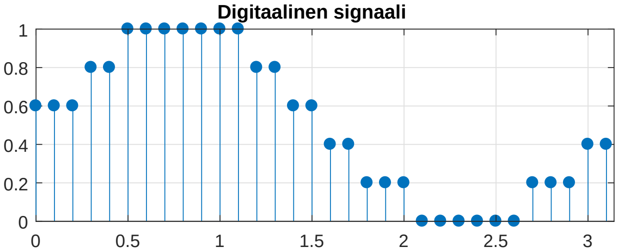
Signaalit ja järjestelmät

- **Diskreettiaikainen signaali** saa arvoja vain tietyillä ajanhetkillä.



Signaalit ja järjestelmät

- **Digitaalinen signaali** saa vain äärellisen määrän eri arvoja ja vain tietyillä ajanhetkillä.



Signaalit ja järjestelmät

- Matemaattisen käsittelyn helpottamiseksi lukujonoja käytetään usein diskreettiaikaisen signaalin mallina.
- Toisin kuin reaalimaailman signaali, lukujono on äärettömän pitkä, mutta tämä ei ole mallinnuksen kannalta ongelma.

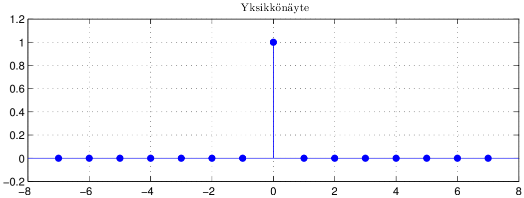
Tampereen yliopisto

Eräitä signaaleita

- **Yksikkönäyte** (unit sample) eli **impulssi** $\delta(n)$ määritellään seuraavasti:

$$\delta(n) = \begin{cases} 1, & \text{kun } n = 0, \\ 0, & \text{kun } n \neq 0. \end{cases}$$

$n = -2$
$n + 2 = 0$

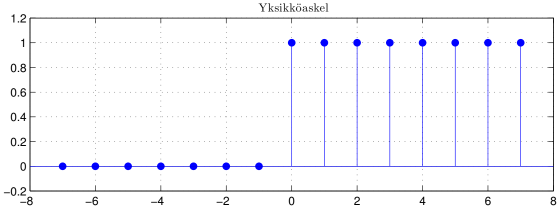


Matlab: `delta = [zeros(1,7),1,zeros(1,7)];`

Eräitä signaaleita

- **Yksikköaskel** (unit step) $u(n)$ määritellään seuraavasti:

$$u(n) = \begin{cases} 1, & \text{kun } n \geqslant 0, \\ 0, & \text{kun } n < 0. \end{cases}$$



Matlab: `u = [zeros(1,7),ones(1,8)];`

Eräitä signaaleita

- Nämä jonot voidaan esittää toistensa avulla seuraavasti:

$$\delta(n) = u(n) - u(n-1)$$

ja

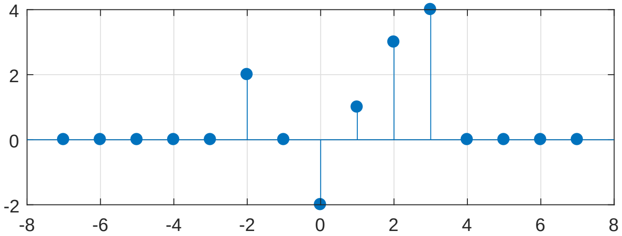
$$u(n) = \sum_{k=-\infty}^{n} \delta(k).$$

- Mikä tahansa jono voidaan esittää siirrettyjen ja painotettujen yksikkönäytteiden avulla seuraavasti:

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

Eräitä signaaleita

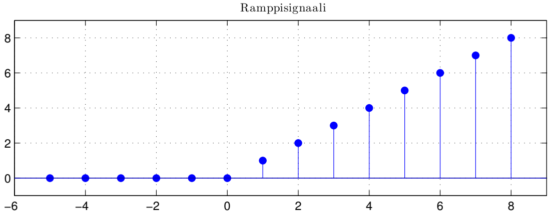
- Kyseinen esitysmuoto on erikoistapaus **konvoluutiosta**, johon tutustutaan kohta lähemmin.
- Esimerkiksi alla olevan kuvan signaali voidaan esittää muodossa $x(n) = 2\delta(n+2) - 2\delta(n) + \delta(n-1) + 3\delta(n-2) + 4\delta(n-3)$.



Eräitä signaaleita

- **Ramppisignaali** (ramp signal) määritellään seuraavasti:

$$r(n) = nu(n) = \begin{cases} n, & \text{kun } n \geqslant 0, \\ 0, & \text{kun } n < 0. \end{cases}$$

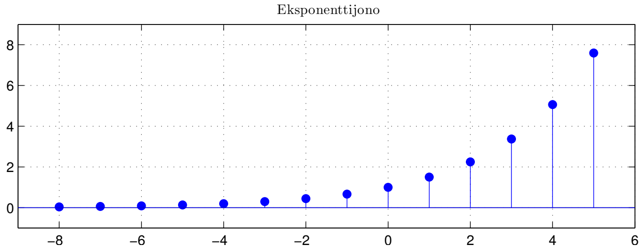


Matlab: `r = [zeros(1,7),0:7];`

Eräitä signaaleita

- **Eksponenttijono** määritellään seuraavasti:

$$x(n) = A\alpha^n.$$



Matlab: `x = -7:7; e = 1.5.^x;`

Jonojen ominaisuuksia

- Diskreettiaikainen signaali on **jaksollinen**, jos on olemassa sellainen $N \in \mathbb{N}$, että

$$x(n) = x(n + N),$$

  kaikilla indeksin n arvoilla.
- Lukua N sanotaan **jakson pituudeksi**.
- Usein taajuuden f (Hz) sijaan käytetään **kulmataajuutta** ω (rad/s).
- Näiden välinen yhteys on $\boxed{\omega = 2\pi f}$.

Jonojen ominaisuuksia

- Esimerkiksi sinijono $x(n) = \sin(\omega n)$ on jaksollinen.
- Sen jakson pituus on
$$N = \frac{2\pi}{\omega} = \frac{1}{f}.$$
- Jono on tosin määritelmän tiukassa mielessä jaksollinen vain jos näin saatu N on kokonaisluku.
- Jos $\mathcal{S}$ on kaikkien signaalien joukko, niin kuvausta $\mathcal{F} : \mathcal{S} \mapsto \mathcal{S}$ sanotaan diskreetiksi järjestelmäksi tai suotimeksi.
- Kuvauksen $\mathcal{F}$ argumenttia kutsutaan **sisäänmenojonoksi** tai **herätteeksi** ja sen palauttamaa jonoa kutsutaan **ulostulojonoksi** tai **vasteeksi**.

Jonojen ominaisuuksia

- Esimerkiksi yhtälö $y(n) = x(n-10)$ määrittelee suotimen, joka viivästää signaalia 10 askelta.
- Toinen esimerkki laskee keskiarvon:

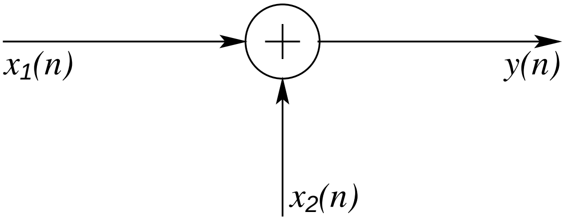
$$y(n) = \frac{1}{5}(x(n) + x(n-1) + x(n-2) + x(n-3) + x(n-4))$$

tai yleisemmin:

$$y(n) = \frac{1}{2K+1}\sum_{k=0}^{2K} x(n-k).$$

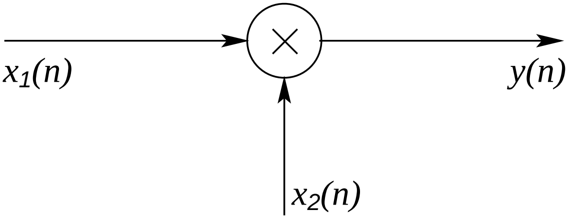
Perusoperaatiot jonoille

- Perusoperaatiot lukujonoille (signaaleille) on esitetty seuraavassa.
- Mukana ovat myös operaatioille lohkokaavioissa käytettävät symbolit.
- **Yhteenlasku:** $y(n) = x_1(n) + x_2(n)$

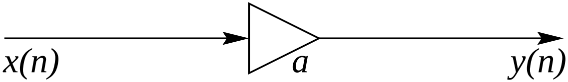


Perusoperaatiot jonoille

- **Kertolasku:** $y(n) = x_1(n) \cdot x_2(n)$

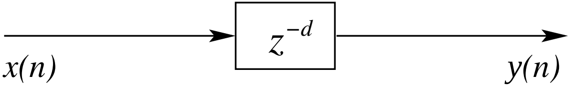


- **Vakiolla kertominen:** $y(n) = ax(n)$



Perusoperaatiot jonoille

- **Viive:** $y(n) = x(n-d)$

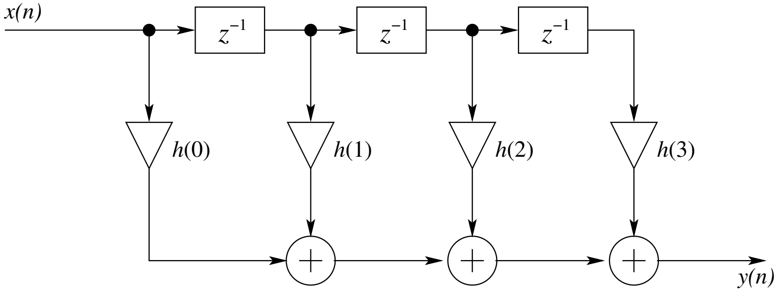


- Edellä olleita operaatioita voidaan yhdistellä ja muodostaa näin monimutkaisempia järjestelmiä.
- Yksi esimerkki on järjestelmä

$$y(n) = \sum_{k=0}^{3} h(k)x(n-k),$$

jonka lohkokaavio on seuraavassa kuvassa.

Perusoperaatiot jonoille



$$y(n) = \sum_{k=0}^{3} h(k)x(n-k)$$