Audiotuning og DFT

Christian Henriksen

8. september 2017

1 Indledning, notation og forberedelse

I dette materiale studerer vi, hvordan en audiotuner virker. En audiotuner kan bruges til at stemme et musikinstrument som for eksempel en guitar. Den kan afgøre om en tone er ren eller falsk og hvis den er falsk, så kan den angive om tonen skal være dybere eller højere for at nå hen til den nærmeste rene tone.

En måde at konstruere en audiotuner er ved at bruge den diskrete fouriertransformation (DFT). Løst sagt går fourieranalyse ud på at opløse et signal i rene svningninger. Den diskrete fouriertransformation kan man omforme rækken af tal til en anden række af tal, der beskriver frekvenser og amplituder af den oprindelige række af tal.

Inden vi kan kaste os over fourieranalysen, har vi dog brug for at definere en del begreber.

1.1 Mængder og talsæt

En mængde er en samling elementer hvor rækkefølgen ikke betyder noget. Et objekt er enten med i en mængde (dvs et element i mængden) eller også er det det ikke. Eksempler på mængder er mængden formet af et og tre: $\{1, 3\} = \{3, 1\}$, de hele tal $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$, de reelle tal $\mathbb{R}$, og de komplekse tal $\mathbb{C}$. At 3 er et element i $\{1, 3\}$ skrives kort $3 \in \{1, 3\}$. En meget vigtig mængde er den tomme mængde $\emptyset$, der ikke indeholder nogen elementer. mængde

Et talsæt er en samling elementer, opskrevet i en bestemt rækkefølge. Fx er $(1, 2)$ og $(2, 1)$ to *forskellige* sæt. Talsættet $(1, 1)$ er ikke det samme som talsættet (1). talsæt

Alle talsæt på formen (x, y), hvor x og y er reelle tal, danner en mængde der betegnes $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$. Så $\mathbb{R}^2$ er en mængde, og et element i $\mathbb{R}^2$ er et talsæt. Lad nu N være et positivt heltal. Mængden af alle talsæt $(x_0, x_1, \dots, x_{N-1})$, hvor $x_j \in \mathbb{R}$ for $j = 0, 1, \dots, N - 1$ betegnes $\mathbb{R}^N$.

Opgave 1 Hvad tror du $\mathbb{C}^2$ betegner, og mere generelt, hvad betegner $\mathbb{C}^N$? Gælder der $(z_0, z_1, \ldots, z_N) \in \mathbb{C}^N$?

Ordet *mængde* hedder *set* på engelsk. Så et engelsk *set* er altså ikke det samme som et dansk *sæt*.

1.2 Periodiske følger

Lad $x : \mathbb{Z} \to \mathbb{C}$ være en afbildning. Det betyder at for alle heltal n betegner $x(n)$ et komplekst tal. Siden x ikke er defineret på kontinuet $\mathbb{R}$, men kun på mængden $\mathbb{Z}$, kalder vi x en dobbelt-uendelig følge, eller bare en følge. Sædvanligvis skriver man x_n i stedet for $x(n)$ for følger. følge

Følgen x kaldes N-periodisk, hvis der findes et positivt heltal $N \in \mathbb{N}$, således at periodisk

$$\text{for alle heltal } n \text{ gælder } x_n = x_{n+N}.$$

Periodicitetsbetingelsen kan skrives mere kompakt ved at bruge alkvantoren $\forall$. Symbolet $\forall$ læses *for alle*, og vi kan skrive betingelsen $\forall$

$$\forall n \in \mathbb{Z} \qquad x_n = x_{n+N}. \tag{1}$$

Du vil snart høre mere om alkvantoren i diskret matematik.

Hvis en følge x er 1-periodisk, følger det af periodicitetsbetingelsen (1), at

$$\cdots = x_{-2} = x_{-1} = x_0 = x_1 = x_2 = \cdots .$$

Følgen kan altså kun antage én værdi, og den er derved entydigt bestemt ved at angive denne værdi.

Opgave 2 Vis at en 2-periodisk følge er entydigt bestemt ved talsættet

$$(x_0, x_1) \in \mathbb{C}^2.$$

Helt generelt gælder at en N-periodisk følge er entydigt bestemt ved talsættet $(x_0, x_1, \ldots, x_{N-1}) \in \mathbb{C}^N$. For at se det indfører vi resten ved division med N.

Egentligt kender vi allerede begrebet *rest* fra 2. klasse. Hvis $N = 7$ elever skal dele 15 nødder, så er der $q = 2$ til hver og $r = 1$ til overs, for der gælder jo $15 = 2 \cdot 7 + 1$. Generelt gælder at alle $n \in \mathbb{Z}$ på entydig måde kan skrives som $n = qN + r$, hvor $q \in \mathbb{Z}$ og $r \in \{0, 1, \ldots, N-1\}$. Man kalder r for *resten* ved division af n med N. I denne note benytter vi betegnelsen $[n]$ for resten ved division af n med N. Hvis $(x_0, x_1, \ldots, X_{N-1}) \in \mathbb{C}^N$, så er $x_{[n]}$ en N-periodisk følge. rest [n]

Eksempel 1

Der gælder N går op i k, hvis og kun hvis $[k] = 0$.

Antag x er en N-periodisk følge. Vi viser at den er entydigt bestemt ved de første N elementer $(x_0, x_1, \dots, x_{N-1}) \in \mathbb{C}^N$. Ved at bruge periodicitetsbetingelsen (1) gentagne gange, ser vi

$$x_n = x_{n+N} = x_{n+2N} = \cdots .$$

Ved at bruge periodicitetsbetingelsen den anden vej, ser vi at der også gælder

$$x_n = x_{n-N} = x_{n-2N} = \cdots .$$

Generelt gælder der derfor

$$\forall j \in \mathbb{Z} \qquad x_n = x_{n+jN}.$$

Da nu $x_n = x_{qN+r} = x_r$, er det nok at kende $(x_0, x_1, \dots, x_{N-1}) \in \mathbb{C}^N$ til at rekonstruere hele følgen x_n.

Omvendt, hvis vi kender hele følgen $\dots, x_{-2}, x_{-1}, x_0, x_1, x_2, \dots$, så kender vi selvfølgelig også $(x_0, x_1, \dots, x_{N-1})$.

Opgave 3 Lad $N = 3$. Hvis $(x_0, x_1, x_2) = (7, 9, 13) \in \mathbb{R}^N$, hvad er da $x_{[10]}$?

1.3 Summer og sumtegn

I denne note har vi tit brug for at skrive summer. Heldigvis findes der en praktisk måde at angive summer på ved hjælp sumtegnet, $\sum$, der er det græske bogstav store sigma. I det følgende forklares sumtegnets betydning, og der postuleres nogle konventioner og regneregler for summer. $\sum$

Vi starter med at se på et simpelt eksempel. Antag at x_1, x_2, x_3 og x_4 er tal. Så er $S = x_1 + x_2 + x_3 + x_4$ en sum af de fire tal. Denne sum kan opskrives ved hjælp af sumtegnet. Sætter vi $I = \{1, 2, 3, 4\}$ forstår man ved symbolet $\sum_{k \in I} x_k$ netop summen S af de fire led. Man kalder k for et indeks og I for indeksmængden, dvs mængden vi summerer over. Et andet eksempel er summen $x_1 + x_3$. Den kan opskrives ved hjælp af sumtegnet, nemlig som $\sum_{k \in \{1,3\}} x_k$. Man har vedtaget at den tomme sum er lig 0: $\sum_{k \in \emptyset} x_k = 0$. tom sum

Tit har man brug for at summere op over en fortløbende følge af hele tal. Vender vi tilbage til det første eksempel, var det tallene fra 1 til 4. I stedet for at opskrive indeksmængden, kan man nøjes med at skrive $\sum_{k=1}^{4} x_k$. Tænk over hvorfor der gælder $\sum_{k \in \{1,3\}} x_k = \sum_{k=0}^{1} x_{2k+1}$.

I overenstemmelse med konventionen om at den tomme sum er lig 0, sætter man

$$\sum_{k=a}^{b} x_k = 0$$

når $a > b$.

Ligesom man ganger før man adderer, har man også en rækkefølge man udfører summationer i. Man udregner sumtegnet efter multiplikation og division, men før addition og subtraktion. Så der gælder

$$\sum_{k\in\{1,2\}} kx_k^k = 1x_1^1 + 2x_2^2,$$

$$\sum_{k\in\{1,2\}} x_k + 3 = x_1 + x_2 + 3.$$

Helt som vi er vant til, kan operationernes rækkefølge ændres ved brug af parenteser.

$$\sum_{k\in\{1,2\}} (x_k + 3) = x_1 + 3 + x_2 + 3.$$

Eftersom man kan sætte ud foran en parentes i en sum, og rækkefølgen af leddene i en sum er ligegyldig, gælder følgende to regneregler.

$$\sum_{k\in I} ax_k = a\sum_{k\in I} x_k$$

$$\sum_{k\in I}(x_k + y_k) = \sum_{k\in I} x_k + \sum_{k\in I} y_k.$$

Der gælder

$$\sum_{k=a}^{b} x_k = \sum_{k=a-c}^{b-c} x_{k+c}. \tag{2}$$

For at indse at dette er sandt, tænk over hvor mange led der er i de to summer, og hvad det første og sidste led er i de to tilfælde.

Hvis x_k er en N-periodisk følge gælder for et vilkårligt heltal j at

$$\sum_{k=0}^{N-1} x_k = \sum_{k=j}^{j+N-1} x_k. \tag{3}$$

for begge summer er foretaget over netop en periode.

Opgave 4 Beregn

$$\sum_{k=1}^{n} k$$

for $n = 0, 1, 2$ og 3. Giver summen det samme som $n(n+1)/2$ i de fire tilfælde?

Opgave 5 Find ud af hvad der skal stå i stedet for $\square$, for at der gælder

$$\sum_{k\in\{0,2,4\}} x_k = \sum_{k=0}^{2} x_{\square}.$$

1.4 Samplede signaler

Mange ting vi interesserer os for kan bedst beskrives ved en funktion $g(t)$. Det kunne være temperaturen i et værelse der varierer med tiden t, eller det kunne være positionen af en membranen i en mikrofon der også varierer med tiden t. Tit og ofte har vi ikke adgang til funktionen $g(t)$, men kun til nogle af funktionsværdierne $x_n = g(t_n)$ for $t = 0, 1, \dots, N-1$. Hvis vi for eksempel måler temperaturen i et værelse 1 gang i timen, over et døgn, har vi kun 24 funktionsværdier at analysere på, ikke de uendelige mange der findes i signalet $g(t)$. Man kalder funktionen $g(t)$ for et signal, mens man kalder funktionsværdierne x_n for et diskretiseret eller samplet signal.

samplet signal

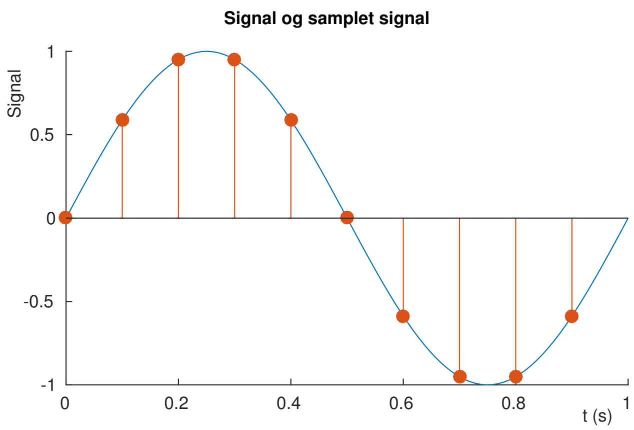
Typisk når man sampler et signal $g(t)$ er tiden mellem hver måling den samme. Det vil sige sammenhængen mellem det diskrete signal x og det tidskontinuerte signal g er

$$x_n = g(t_0 + n \cdot t_s)$$

Tidsrummet t_s, som altså er tiden mellem hver måling (eller sampling), kaldes *sampletiden*. Den reciprokke $f_s = \frac{1}{t_s}$ kaldes *samplefrekvensen*. Samplefrekvensen udtrykker hvor mange samplinger vi foretager per tidsenhed.

sampletid

samplefrekvens



Opgave 6 På grafen kan man se et signal, her en ren svingning, og det tilhørende samplede signal. Angiv sampletiden t_s og samplefrekvensen f_s.

2 Svingninger og lyd

Når periodiske trykbølger får luften til at vibrere indenfor et vist frekvensområde, så kan det menneskelige øre opfange signalet, og vi hører lyd. Lyd og svingninger hører altså sammen. Vi kigger derfor først på en særlig simpel type periodisk funktion, som vi kalder en ren svingning.

Når lufttrykket svinger med en bestemt frekvens opfatter vi det som sagt som lyd. Indenfor musikken interesserer man sig selvfølgelig for lyd. Nogle lyde kalder man for rene toner, andre for falske. Der er lidt forskellige systemer, men det mest anvendte i den vestlige verden er et tonesystem der kaldes *den ligesvævende stemning*. Den giver vi en kort introduktion til i næste afsnit, men først ser vi på rene svingninger.

2.1 Ren svingning

En ren svingning også nogle gange kaldet en harmonisk svingning er en bestemt slags periodisk funktion. Den præcise definition er som følger.

Definition 1
En *ren svingning* er en funktion $g(t)$, der kan skrives på formen

$$g(t) = A\sin(2\pi f t + \phi),$$

for konstanter $A, f, \phi \in \mathbb{R}$. Konstanten A kaldes *amplituden*, konstanten f for *frekvensen* og konstanten ϕ for *faseforskydningen*.

Vi kan også skrive

$$g(t) = A\sin(\omega t + \phi),$$

hvor $\omega = 2\pi f$ kaldes *vinkelfrekvensen*.

Opgave 7 Se på funktionen $g(t) = \sin(2\pi t)$.

(a) Tegn grafen ved hjælp af et program. I Matlab kan det gøres ved hjælp af følgende to linjer

```
X = linspace (-0.5, 1.5, 101);
plot (X, sin (2.0 * pi * X));
```

(b) Prøv at ændre amplituden. Hvad sker der geometrisk med grafen? Hvordan kan amplituden aflæses på grafen?

(c) Prøv at ændre på frekvensen. Hvad sker der geometrisk på grafen? Størrelsen $T = \frac{1}{f}$ er perioden for g. Den beskrives hvor mange svingninger der sker på en tidsenhed. Hvordan kan den aflæses på grafen?

(d) Prøv slutteligt at ændre på faseforskydningen. Hvad sker der geometrisk på grafen? Hvad sker der hvis faseforskydningen ændres med mere end 2π?

2.2 Den ligesvævende stemning

I den vestlige musik benytter man almindeligvis den ligesvævende stemning. I denne har man en kammertone der svinger med en bestemt grundfrekvens f_g. Der er lidt variation i hvad man bruger som grundfrekvens, men et typisk tal er $f_g = 440\,\mathrm{Hz}$, og det holder vi os til i denne note. En ren tone svinger med en frekvens ν der kan skrives

grund-frekvens

$$\nu = f_g \cdot 2^{s/12}, \tag{4}$$

hvor $s \in \mathbb{Z}$ er et heltal. For eksempel kan man finde frekvensen for tonen lige over kammertonen, ved at man sætter $s = 1$. Så får man

$$\nu = f_g \cdot 2^{1/12} \approx 466.16\,\mathrm{Hz}.$$

Når man går fra en værdi af k til den næste, siger man at man går en halvtone op. Tonen svarende til frekvensen 466.16 Hz ligger altså en halvtone over kammertonen. Et spænd på 12 halvtoner kaldes en oktav.

halv-tone

Opgave 8 Vis at det at gå en oktav op svarer til en fordobling af frekvensen.

Hvis vi kender frekvensen ν, kan vi finde den tilhørende s ved hjælp af ligning (4). Isolerer vi s får vi

$$s(\nu) = 12 \cdot \frac{\log(\nu/f_g)}{\log(2)}.$$

Så hvis s er et heltal, har vi at gøre med en ren tone. Hvis s ikke er et heltal har vi at gøre med en falsk tone, jo længere væk fra et heltal s er, jo mere falsk klinger den.

Opgave 9 Svarer frekvensen $\nu = 520\,\mathrm{Hz}$ til en ren tone? Hvad med $\nu = 220\,\mathrm{Hz}$?

Vi kunne angive foreskellen mellem to frekvenser ν_1 og ν_2 ved $s(\nu_2) - s(\nu_1)$, men man bruger i stedet en finere skala, man måler nemlig forskellen i cent, der er 100 gange forskellen $s(\nu_2) - s(\nu_1)$.

cent

Opgave 10 Vis at der fra ν_1 til ν_2 er

$$1200\,\mathrm{cent} \cdot \frac{\log\left(\frac{\nu_2}{\nu_1}\right)}{\log(2)}.$$

Hermed har vi set på det mindst mulige teori om frekvenser og toner til at kunne konstruere en audiotuner. En mere omfattende fremstilling kan findes på `www.musikpedia.dk/lyd`.

2.3 Den komplekse eksponentialfunktion

Den komplekse eksponetialfunktion $z \mapsto \mathrm{e}^z$ er en afbildning fra $\mathbb{C}$ til $\mathbb{C}$. Det betyder at givet et $z \in \mathbb{C}$, tilskriver funktionen et tal $w = \mathrm{e}^z$ i $\mathbb{C}$. Den komplekse eksponentialfunktion er defineret ved hjælp af de reelle funktioner, e^x, sin og cos. Skriver vi nemlig $z = x + \mathrm{i}y$, hvor x og y er reelle, definerer man e^z

$$\mathrm{e}^z = \mathrm{e}^x (\cos y + \mathrm{i} \sin y) . \tag{5}$$

Den komplekse eksponentialfunktion opfylder følgende fundamentale ligning.

$$\mathrm{e}^{z_1+z_2} = \mathrm{e}^{z_1}\mathrm{e}^{z_2}. \tag{6}$$

Denne ligning som vi vil få brug for igen og igen, kan bevises ved hjælp af defintionen og additionsformlerne.

Opgave 11 Bevis at ligning (6) gælder. Gå fx frem som følger. Skriv

$$z_1 = x_1 + \mathrm{i}y_1, \quad z_2 = x_2 + \mathrm{i}y_2.$$

Der gælder

$$\begin{aligned} \mathrm{e}^{z_1+z_2} &= \mathrm{e}^{x_1+\mathrm{i}y_1+x_2+\mathrm{i}y_2} \\ &= \mathrm{e}^{x_1+x_2+\mathrm{i}(y_1+y_2)} \\ &= \mathrm{e}^{x_1+x_2} \left(\cos(y_1 + y_2) + \mathrm{i} \sin(y_1 + y_2)\right) . \end{aligned}$$

Brug nu at du ved at (6) gælder for reelle tal, samt additionsformlerne for cosinus og sinus til at nå frem til højresiden i (6). Det kan være en hjælp at prøve at arbejde baglæns fra højresiden. Vink: Additionsformlerne er

$$\begin{aligned} \cos(x + y) &= \cos(x)\cos(y) - \sin(x)\sin(y) \\ \sin(x + y) &= \sin(x)\cos(y) + \cos(x)\sin(y). \end{aligned}$$

Sinusfunktionen er hvad man kalder en ulige funktion. Der gælder at $\sin(-y) = -\sin(y)$. Cosinusfunktionen er en lige funktion for der gælder $\cos(-y) = \cos(y)$. Det følger at ulige lige

$$\mathrm{e}^{\mathrm{i}y} + \mathrm{e}^{-\mathrm{i}y} = 2\cos y,$$

for vi har

$$\begin{aligned} \mathrm{e}^{\mathrm{i}y} + \mathrm{e}^{-\mathrm{i}y} &= \cos(y) + \mathrm{i}\sin(y) + \cos(-y) + \mathrm{i}\sin(-y) \\ &= \cos(y) + \cos(y) + \mathrm{i}(\sin(y) - \sin(y)) \\ &= 2\cos(y). \end{aligned}$$

Opgave 12 Bevis at der gælder

$$e^{iy} - e^{-iy} = 2i \sin y.$$

Sammenlagt gælder altså

$$\cos y = \frac{1}{2}\left(e^{iy} + e^{-iy}\right) \tag{7}$$

$$\sin y = \frac{1}{2i}\left(e^{iy} - e^{-iy}\right) \tag{8}$$

Eksponentialfunktionen indeholder på en gang information om sinus og cosinus, og den er mange gange nemmere at have gøre med end sinus og cosinus. Når man vil udlede en formel med sinus og cosinus kan det nogle gange svare sig at skrive om til udtryk med eksponentialfunktionen og udlede formlen ved hjælp af denne.

2.4 Eksponentialfunktionen og rene svingninger

Vi har defineret en ren svingning ved hjælp af sinus, se Definition 1, men vi kunne lige så vel have brugt cosinusfunktion, for der gælder at

$$\cos(x) = \sin(x + \frac{\pi}{2}).$$

Derfor er

$$A \sin(2\pi f t + \phi) = A \cos(2\pi f t + \phi - \frac{\pi}{2}).$$

Hvis man tager realdelen (eller imaginærdelen) af den komplekse funktion $Ce^{i\omega t}$, hvor $C \in \mathbb{C}$ er en konstant, så får man en ren svingning, og omvendt kan en ren svingning altid skrives som realdelen af $Ce^{i\omega t}$. Det undersøger vi i følgende opgave.

Opgave 13 Vis at

$$\mathrm{Re}(Ce^{i\omega t}) = A \cos(\omega t + \phi),$$

når

$$C = Ae^{i\phi}.$$

3 Diskret fouriertransformation

En mikrofon måler lydtrykket som funktion af tiden. Hvis man sampler lydtrykket, så får man et diskret signal. For at lave en audiotuner, har vi brug for analysere det diskrete signal og forstå det som toner, eller mere præcist, forstå hvilke frekvenser der indgår i signaler. Så kan vi nemlig afgøre om lyden mikrofonen har opfanget er ren eller ej. Den matematisk afbildning der går fra et samplet signal og giver os frekvenser og amplituder, er den diskrete fouriertransformation. På nuværende tidspunkt har vi indført nok notation til at kunne indføre den diskrete fouriertransformation.

3.1 Indføring af diskret fouriertransformation

Definitionen på den diskrete fouriertransformation er som følger. DFT

Definition 2
Givet $z = (z_0, z_1, \dots, z_{N-1}) \in \mathbb{C}^N$, definerer vi $Z \in \mathbb{C}^N$ ved

$$Z_k = \sum_{n=0}^{N-1} z_n \mathrm{e}^{-\frac{2\mathrm{i}\pi kn}{N}}, \quad k = 0, 1, \dots, N-1.$$

Afbildningen $z \mapsto Z$ kaldes den diskrete fouriertransformation, og betegnes med DFT.

Læg mærke til at højresiden i ligningen i den foregående definition giver mening for alle $k \in \mathbb{Z}$. Tillader vi $k \in \mathbb{Z}$, bliver Z en N-periodisk følge, og $z_{[k]} = z_k$ for alle $k \in \mathbb{Z}$.

Vi definerer også IDFT. IDFT

Definition 3
Givet $Z = (Z_0, Z_1, \dots, Z_{N-1}) \in \mathbb{C}^N$, definerer vi $z = (z_0, z_1, \dots, z_{N-1}) \in \mathbb{C}^N$ ved

$$z_n = \frac{1}{N} \sum_{k=0}^{N-1} Z_k \mathrm{e}^{\frac{2\mathrm{i}\pi kn}{N}}, \quad n = 0, 1, \dots N-1.$$

Afbildningen $Z \mapsto z$ kaldes den inverse diskrete fouriertransformation, og betegnes med IDFT.

Ligesom det var tilfældet med definitionen af DFT giver højresiden i definitionens ligning mening for alle $n \in \mathbb{Z}$. Bruger vi formlen for et vilkårligt $n \in \mathbb{Z}$ gælder $z_n = z_{[n]}$, og z kan altså opfattes som en N-periodisk følge.

Eksempel 2
Hvis $N = 2$ og $(z_0, z_1) \in \mathbb{C}^2$, og vi sætter $Z = \mathrm{DFT}(z_0, z_1)$, så er

$$\begin{aligned} Z_0 &= z_0 \mathrm{e}^0 + z_1 \mathrm{e}^0 = z_0 + z_1 \\ Z_1 &= z_0 \mathrm{e}^0 + z_1 \mathrm{e}^{-\mathrm{i}\pi} = z_0 - z_1 \end{aligned}$$

Hvis på den anden side $N = 2$, $(Z_0, Z_1) \in \mathbb{C}^2$, og vi sætter $z = \mathrm{IDFT}(Z_0, Z_1)$, så er

$$\begin{aligned} z_0 &= \frac{1}{2}(Z_0 \mathrm{e}^0 + Z_1 \mathrm{e}^0) = \frac{1}{2}(Z_0 + Z_1) \\ z_1 &= \frac{1}{2}(Z_1 \mathrm{e}^0 + Z_1 \mathrm{e}^{\mathrm{i}\pi}) = \frac{1}{2}(Z_0 - Z_1) \end{aligned}$$

Opgave 14 Sæt $\omega = \mathrm{e}^{2\mathrm{i}\pi/3} = -\frac{1}{2} + \mathrm{i}\frac{\sqrt{3}}{2}$. Lad $(z_0, z_1, z_2) \in \mathbb{C}^3$ være givet, og sæt $(Z_0, Z_1, Z_2) = \mathrm{DFT}(z_0, z_1, z_2)$. Vis at

$$\begin{aligned} Z_0 &= z_0 + z_1 + z_2 \\ Z_1 &= z_0 + \omega^2 z_1 + \omega z_2 \\ Z_2 &= z_0 + \omega z_1 + \omega^2 z_2. \end{aligned}$$

Lad nu $(Z_0, Z_1, Z_2) \in \mathbb{C}^3$ være givet, og sæt

$$(z_0, z_1, z_2) = \mathrm{IDFT}(Z_0, Z_1, Z_2).$$

Vis at

$$\begin{aligned} z_0 &= \frac{1}{3}(Z_0 + Z_1 + Z_2) \\ z_1 &= \frac{1}{3}(Z_0 + \omega Z_1 + \omega^2 Z_2) \\ z_2 &= \frac{1}{3}(Z_0 + \omega^2 Z_1 + \omega Z_2). \end{aligned}$$

I definitionen på IDFT tager vi summen for $k = 0$ til $N - 1$, men da eksponentialfunktionen er periodisk har vi flere muligheder.

Sætning 1
Hvis $(z_0, \dots, z_{N-1}) = \mathrm{IDFT}(Z_0, \dots, Z_{N-1})$ gælder for alle $a \in \mathbb{Z}$

$$z_n = \frac{1}{N} \sum_{k=a}^{a+N-1} Z_{[k]} \mathrm{e}^{\frac{2\mathrm{i}\pi kn}{N}}, \quad n = 0, 1, \dots N - 1.$$

Specielt gælder, at når N er ulige og altså kan skrives $N = 2M + 1$, at

$$z_n = \frac{1}{N} \sum_{k=-M}^{M} Z_{[k]} \mathrm{e}^{\frac{2\mathrm{i}\pi kn}{N}}, \quad n = 0, 1, \dots N - 1.$$

Når man ser på navnet invers diskret fouriertransformation, kunne man godt få en snigende fornemmelse af, at DFT og IDFT er hinandens inverse. Og det er de: invers

Sætning 2
Lad $(z_0, z_1, \dots, z_{N-1}) \in \mathbb{C}^N$ og $(Z_0, Z_1, \dots, Z_{N-1}) \in \mathbb{C}^N$ være vilkårlige. Der gælder

$$(Z_0, Z_1, \dots, Z_{N-1}) = \mathrm{DFT}(z_0, z_1, \dots, z_{N-1}) \qquad \Leftrightarrow$$
$$(z_0, z_1, \dots, z_{N-1}) = \mathrm{IDFT}(Z_0, Z_1, \dots, Z_{N-1}).$$

Vi beviser først sætningen i tilfældet $N = 2$. Først viser vi at der gælder $\Rightarrow$.

Antag $(Z_0, Z_1) = \mathrm{DFT}(z_0, z_1)$. Ifølge Eksempel 2 er $Z_0 = z_0 + z_1$ og $Z_1 = z_0 - z_1$. Vi sætter $(w_0, w_1) = \mathrm{IDFT}(Z_0, Z_1)$ og skal vise at $w_0 = z_0$ og $w_1 = z_1$. Ved at bruge Eksempel 2 igen, ser vi

$$\begin{aligned} w_0 &= \frac{1}{2}(Z_0 + Z_1) = \frac{1}{2}(z_0 + z_1 + z_0 - z_1) = z_0, \\ w_1 &= \frac{1}{2}(Z_0 - Z_1) = \frac{1}{2}(z_0 + z_1 - (z_0 - z_1)) = z_1. \end{aligned}$$

Hermed har vi vist $\Rightarrow$.

Antag nu omvendt at $(z_0, z_1) = \mathrm{IDFT}(Z_0, Z_1)$, altså $z_0 = \frac{1}{2}(Z_0 + Z_1)$ og $z_1 = \frac{1}{2}(Z_0 - Z_1)$. Sætter vi $(W_0, W_1) = \mathrm{DFT}(z_0, z_1)$, skal vi vise $W_0 = Z_0$ og $W_1 = Z_1$. Vi har

$$\begin{aligned} W_0 &= z_0 + z_1 = \frac{1}{2}(Z_0 + Z_1) + \frac{1}{2}(Z_0 - Z_1) = Z_0, \\ W_1 &= z_0 - z_1 = \frac{1}{2}(Z_0 + Z_1) - \frac{1}{2}(Z_0 - Z_1) = Z_1. \end{aligned}$$

Dette viser at der også gælder $\Leftarrow$, og afslutter beviset i tilfældet $N = 2$.

Opgave 15 Gå frem som ovenfor og brug Opgave 14 til at bevis sætningen i tilfældet $N = 3$.

For at bevise Sætning 2 generelt observerer vi følgende. Hvis $a \in \mathbb{Z}$ opfylder $-N < a < N$, da gælder

$$\sum_{k=0}^{N-1} \mathrm{e}^{\frac{2i\pi ka}{N}} = \begin{cases} N & \text{for } a = 0 \\ 0 & \text{for } a \neq 0 \end{cases} \tag{9}$$

Der er to tilfælde, $a = 0$ og $a \neq 0$. Vi starter med at se på det første. Hvis $a = 0$, så er $\mathrm{e}^{\frac{2i\pi ka}{N}} = 1$, uanset hvad k er. Summen er altså en sum af N ettaller, og giver derfor N.

Tilbage står det generelle tilfælde hvor $a \neq 0$. Vi sætter $S = \sum_{k=0}^{N-1} \mathrm{e}^{\frac{2\mathrm{i}\pi ka}{N}}$ og skal vise at $S = 0$. Ganger vi med $\mathrm{e}^{\frac{2\mathrm{i}\pi a}{N}}$, får vi

$$\begin{aligned}
\mathrm{e}^{\frac{2\mathrm{i}\pi a}{N}} S &= \sum_{k=0}^{N-1} \mathrm{e}^{\frac{2\mathrm{i}\pi(k+1)a}{N}} \\
&= \sum_{k=1}^{N} \mathrm{e}^{\frac{2\mathrm{i}\pi ka}{N}} && \text{ifølge (2)} \\
&= \sum_{k=0}^{N-1} \mathrm{e}^{\frac{2\mathrm{i}\pi ka}{N}} && \text{ifølge (3)} \\
&= S.
\end{aligned}$$

Det følger at

$$\left(\mathrm{e}^{\frac{2\mathrm{i}\pi a}{N}} - 1\right) S = 0,$$

og da $-N < a < N$ og $a \neq 0$, er $\mathrm{e}^{\frac{2\mathrm{i}\pi a}{N}} - 1 \neq 0$. Derfor må $S = 0$. Hermed har vi bevist vores observation.

Vi er nu klar til at bevise Sætning 2. Beviset forløber helt som i specialtilfældet $N = 2$. Udregninger bliver bare lidt mere komplicerede, men hovedparten af det ekstra arbejde har vi allerede udført med ligning (9).

Vi starter med at vise „⇒“. Antag

$$(Z_0, Z_1, \dots, Z_{N-1}) = \mathrm{DFT}(z_0, z_1, \dots, z_{N-1}),$$

hvilket vil sige

$$Z_k = \sum_{n=0}^{N-1} \mathrm{e}^{-\frac{2\mathrm{i}\pi kn}{N}} z_n, \quad k = 0, 1, \dots, N-1.$$

Vi sætter $(w_0, w_1, \dots, w_{N-1}) = \mathrm{IDFT}(Z_0, Z_1, \dots, Z_{N-1})$ og skal vise at $w_n = z_n$ for $n = 0, 1, \dots, N-1$. Der gælder

$$\begin{aligned}
w_n &= \frac{1}{N} \sum_{k=0}^{N-1} \mathrm{e}^{\frac{2\mathrm{i}\pi kn}{N}} Z_k \\
&= \frac{1}{N} \sum_{k=0}^{N-1} \mathrm{e}^{\frac{2\mathrm{i}\pi kn}{N}} \sum_{j=0}^{N-1} \mathrm{e}^{-\frac{2\mathrm{i}\pi jk}{N}} z_j \\
&= \frac{1}{N} \sum_{j=0}^{N-1} \left(\sum_{k=0}^{N-1} \mathrm{e}^{\frac{2\mathrm{i}\pi(n-j)k}{N}} \right) z_j.
\end{aligned}$$

Den inderste sum er på samme form som summen i ligning (9), med $a = n - j$. Det følger at den inderste sum giver nul undtagen når $n = j$, hvor den giver N. Så leddene i den yderste sum er nul, bortset fra det ene led hvor $j = n$, hvor leddet er $Nz_j = Nz_n$. Derfor er

$$w_n = \frac{1}{N} N z_n = z_n,$$

hvilket var det vi skulle vise.

Vi mangler at vise at der gælder „$\Leftarrow$" i Sætning 2. Antag derfor

$$(z_0, z_1, \dots, z_{N-1}) = \mathrm{IDFT}(Z_0, Z_1, \dots, Z_{N-1}),$$

hvilket vil sige

$$z_n = \frac{1}{N} \sum_{k=0}^{N-1} \mathrm{e}^{\frac{2\mathrm{i}\pi kn}{N}} Z_k, \quad n = 0, 1, \dots N - 1.$$

Vi sætter $(W_0, W_1, \dots, W_{N-1}) = \mathrm{DFT}(z_0, z_1, \dots, z_{N-1})$ og skal vise at $W_k = Z_k$ for $k = 0, 1, \dots, N - 1$. Der gælder

$$\begin{aligned} W_k &= \sum_{n=0}^{N-1} \mathrm{e}^{\frac{-2\mathrm{i}\pi kn}{N}} z_n \\ &= \sum_{n=0}^{N-1} \mathrm{e}^{\frac{-2\mathrm{i}\pi kn}{N}} \frac{1}{N} \sum_{j=0}^{N-1} \mathrm{e}^{\frac{2\mathrm{i}\pi jn}{N}} Z_j \\ &= \frac{1}{N} \sum_{j=0}^{N-1} \left(\sum_{n=0}^{N-1} \mathrm{e}^{\frac{2\mathrm{i}\pi(j-k)n}{N}} \right) Z_j. \end{aligned}$$

Den inderste sum er på samme form som summen i ligning (9), med $a = j - k$. Det følger at den inderste sum giver nul undtagen når $k = j$, hvor den giver N. Så leddene i den yderste sum er nul, bortset fra det ene led hvor $j = k$, hvor leddet er $NZ_j = NZ_k$. Derfor er

$$W_k = \frac{1}{N} N Z_k = Z_k,$$

hvilket afslutter beviset.

Indtil videre har vi set på komplekse talsæt, men i mange anvendelser er det reelle talsæt vi er interesserede i. Et talsæt med N reelle tal indeholder på en måde halvt så mange værdier som et talsæt med N komplekse tal. Vi kan rekonstruere talsættet $(x_0, \dots, x_{N-1})$ ud fra viden om $(Z_0, \dots, Z_{N-1}) = \mathrm{DFT}(x_0, \dots, x_{N-1})$, så de fouriertransformerede $(Z_0, \dots, Z_{N-1})$ af et reelt talsæt repræsenterer information svarende til N reelle tal. Det giver sig udslag i at der gælder nogle relationer mellem værdierne. Mere præcist har vi følgende sætning. reelle talsæt

Sætning 3
Lad $(Z_0, \dots, Z_{N-1}) = \mathrm{DFT}(z_0, \dots, z_{N-1})$. Der gælder

$$(z_0, \dots, z_{N-1}) \in \mathbb{R}^N$$

hvis og kun hvis

$$Z_{[-k]} = \bar{Z}_k, \qquad k = 0, 1, \dots, N-1.$$

Her har vi brugt notationen $\bar{Z}$ for den kompleks konjugerede af Z. Så hvis $Z = X + \mathrm{i}Y$ så er $\bar{Z} = X - \mathrm{i}Y$.

Læg i øvrigt mærke til, at når $0 < k < N$, er $Z_{[-k]} = Z_{N-k}$. Der gælder $Z_{[0]} = Z_0$.

I tilfældet $N = 3$, medfører sætningen at når $(x_0, x_1, x_2) \in \mathbb{R}^3$, så opfylder $(Z_0, Z_1, Z_2) = \mathrm{DFT}(x_0, x_1, x_2)$ følgende ligninger.

$$\begin{aligned} Z_0 &= \bar{Z}_0 \\ Z_2 &= Z_{[-1]} = \bar{Z}_1 \\ Z_1 &= Z_{[-2]} = \bar{Z}_2. \end{aligned}$$

Den første ligning er ækvivalent med at Z_0 er reel, mens de to næste ligninger er ækvivalente og betyder at Z_1 og Z_2 er hinandens kompleks konjugerede.

Opgave 16 Bevis at for $(x_0, \dots, x_{N-1}) \in \mathbb{R}^N$ gælder $Z_0 \in \mathbb{R}$ på to måder. Først ved at bruge definitionen på DFT; dernæst ved at bruge Sætning 3.

Vi kan bevise Sætning 3 ved direkte udregninger. Antag først at

$$(z_0, \dots, z_{N-1}) \in \mathbb{R}^N.$$

Da

$$\mathrm{e}^{-\frac{2\mathrm{i}\pi(-k)n}{N}} = \mathrm{e}^{\frac{2\mathrm{i}\pi kn}{N}} = \overline{\mathrm{e}^{-\frac{2\mathrm{i}\pi kn}{N}}}$$

gælder

$$\begin{aligned} Z_{[-k]} &= \sum_{n=0}^{N-1} \mathrm{e}^{-\frac{2\mathrm{i}\pi(-k)n}{N}} z_n && \text{og da } z_n \text{ er reel} \\ &= \overline{\sum_{n=0}^{N-1} \mathrm{e}^{-\frac{2\mathrm{i}\pi kn}{N}} z_n} \\ &= \bar{Z}_k. \end{aligned}$$

Antag omvendt at $Z_{[-k]} = \bar{Z}_k$. Ifølge Sætning 2 er

$$(z_0, \dots, z_{N-1}) = \mathrm{IDFT}(Z_0, \dots, Z_{N-1}).$$

Specielt er

$$\begin{aligned}
z_n &= \frac{1}{N} \sum_{k=0}^{N-1} \mathrm{e}^{\frac{2\mathrm{i}\pi kn}{N}} Z_k \\
&= \frac{1}{2N} \left(\sum_{k=0}^{N-1} \mathrm{e}^{\frac{2\mathrm{i}\pi kn}{N}} Z_k + \sum_{k=-(N-1)}^{0} \mathrm{e}^{\frac{2\mathrm{i}\pi kn}{N}} Z_{[k]} \right) \\
&= \frac{1}{2N} \sum_{k=0}^{N-1} \left(\mathrm{e}^{\frac{2\mathrm{i}\pi kn}{N}} Z_k + \mathrm{e}^{\frac{-2\mathrm{i}\pi kn}{N}} Z_{[-k]} \right) \\
&= \frac{1}{2N} \sum_{k=0}^{N-1} \left(\mathrm{e}^{\frac{2\mathrm{i}\pi kn}{N}} Z_k + \mathrm{e}^{\frac{-2\mathrm{i}\pi kn}{N}} \bar{Z}_k \right) \\
&= \frac{1}{N} \sum_{k=0}^{N-1} \mathrm{Re}(\mathrm{e}^{\frac{2\mathrm{i}\pi kn}{N}} Z_k).
\end{aligned}$$

I tredje linje brugte vi at $Z_{[k]}$ er N-periodisk og (3). I sidste linje brugte vi, at for et vilkårligt komplekst tal W gælder $W + \bar{W} = 2\,\mathrm{Re}(W)$. Det ses af udtrykket at z_n er reel, hvilket afslutter beviset af sætningen.

Vi slutter afsnittet af med at nævne nogle yderligere egenskaber ved DFT.

Sætning 4

- Lad $z \in \mathbb{C}^N$ og $w \in \mathbb{C}^n$. Så er

$$\mathrm{DFT}(z + w) = \mathrm{DFT}(z) + \mathrm{DFT}(w).$$

- Lad $z \in \mathbb{C}^N$ og $k \in \mathbb{C}$. Så gælder

$$\mathrm{DFT}(kz) = k\,\mathrm{DFT}(z).$$

4 Afleveringsopgave – audiotuner

En audiotuner kan bruges til at stemme et instrument eller tjekke om man synger rent. Det foregår ved at lyden analyseres, hvorefter den dominante tone angives sammen med afvigelsen fra en ren tone. I dette afsnit vil vi se på, hvordan matematikken i en audiotuner kan se ud.

Afleveringsopgaven går ud på at beskrive matematikken bag en audiotuner. Vi lægger ud med en opvarmningsopgave, der viser at hvis vi kender frekvensen af en svingning, så kan vi afgøre om tonen er ren, eller om frekvensen er for høj eller lav.

Afleveringsopgave 1 Givet en frekvens ν, argumentér for at den nærmeste rene tone højst er 50 cent væk. Angiv en formel for den nærmeste rene tone til ν, og hvor meget den aktuelle tone afviger fra denne. Den nærmeste rene tone angives i halvtoner fra kammertonen, mens afvigelse fra denne angives i cent.

Vink: Det er tilladt og kan være praktisk at bruge funktionen round der runder af tilnærmeste hele tal, så der for eksempel gælder at round(11,2) = 11 og round(11,7) = 12.

Så givet en lydoptagelse, skal vi finde den dominerende frekvens, og afgøre om det svarer til en ren tone eller ej. Men typisk er vi ikke interesseret i hele signalet, kun i et bestemt tidsinterval, for hvilken tone der spilles (eller forsøges at spille) kan jo afhænge af tid. Vi antager at vi nøjes med et se på et signal $g(t)$ i et begrænset tidsinterval, som man kalder et *vindue*. Lad os for at gøre det enkelt, antage at det samplede lyd signal $x_n = g(nt_s), n = 0, 1, \dots, N-1$ svarer til tidsvinduet, og at $N = 2M + 1$ er ulige. vindue

Lad $(Z_0, Z_1, \dots, Z_{N-1}) = \mathrm{DFT}(x_0, x_1, \dots, x_{N-1})$.

Afleveringsopgave 2 Forklar hvorfor der gælder

$$x_n = \frac{1}{N} \sum_{k=-M}^{M} Z_{[k]} \mathrm{e}^{2\mathrm{i}\pi kn/N}.$$

Afleveringsopgave 3 Vis at når $t = nt_s = n/f_s$ for et $n \in \{0, 1, \dots, N-1\}$, gælder

$$g(t) = \frac{1}{N} \sum_{k=-M}^{M} Z_{[k]} \mathrm{e}^{2\mathrm{i}\pi f_s kt/N},$$

og hvis vi derfor sætter

$$h(t) = \frac{1}{N} \sum_{k=-M}^{M} \mathrm{Re}\left(Z_{[k]} \mathrm{e}^{2\mathrm{i}\pi f_s kt/N} \right), \text{ for } t \in \mathbb{R},$$

gælder

$$g(t) = h(t)$$

når $t = nt_s$.

Afleveringsopgave 4 Lad $t_s = 1$ og sæt $(x_0, x_1, x_2) = (0, 2, 2)$. Tegn grafen af $h(t)$ fra sidste opgave for $0 \le t \le 2$.

Afleveringsopgave 5 Forklar hvorfor formlen

$$h(t) = \frac{1}{N} \sum_{k=-M}^{M} \mathrm{Re}\left(Z_{[k]} e^{2i\pi f_s kt/N}\right), \text{ for } t \in \mathbb{R},$$

svarer til at opløse i signalet i rene svingninger, med frekvenser $\frac{k}{N} f_s$, og tilhørende amplituder $\frac{|Z_{[k]}|}{N}$, $k = -M, -M+1, \ldots, M-1, M$.

Hvis man tegner grafen for amplituden $\frac{|Z_{[k]}|}{N}$ som funktion af frekvensen $\frac{k}{N} f_s$ for $k = -M, \ldots, M$ får man det man kalder et *amplitudespektrum*. amplitude-spektrum

Afleveringsopgave 6 Se på eksemplet givet i Afleveringsopgave 4. Hvordan ser amplitudespektret ud?

Afleveringsopgave 7 Forklar hvorfor et reelt signal altid giver anledning til at der er symmetri i amplitudespektret.

På grund af symmetrien i amplitudespektret, nøjes man mange gange med at illustrere amplituden som funktion af de ikke-negative frekvenser.

Ved at have løst Afleveringsopgaverne der er stillet indtil nu, har vi forstået hvad der skal til for at konstruere en tuner. I fildelingen i DTU Inside ligger et matlab script, `audiotune.m`, der laver et spektrogram og finder den dominerende frekvens. Der ligger også to lydfiler `Sine_440.wav` og `floejte.wav`, vi kan teste scriptet på. Den første er en syntetisk genereret lydfil, og den anden er en optagelse af en blokfløjtetone. Men vi mangler at implementere en ting i scriptet, nemlig at beregne afvigelsen fra en ren tone i cent.

Afleveringsopgave 8 Brug jeres viden fra Afleveringsopgave 1 til at få scriptet til at virke. Svarer det syntetiske signal til en ren tone? Hvad med det optagne signal? I kan eksperimentere med at ændre vindueslængden og positionen, og også bruge andre lydfiler, hvis I skulle have lyst.