

Prosjektoppgave i FYS2130 våren 2014

Koherenstid

Dette skrivet består av følgende deler: Innledning, Del 1-5, Referanser og Vedlegg. Du kan jobbe deg gjennom oppgaven på følgende måte:

- Les innledning.
- Les Del 1 og gjennomfør denne delen av prosjektet full ut.
- Les Del 2 og gjennomfør denne, ta så etter tur Del 3, 4 og 5.
- Dersom du ikke rekker Del 4 (evt heller ikke Del 3), bør likevel Del 5 besvares.

Det er selvfølgelig en fordel å skimme gjennom hele skrivet før du starter, men delene er organisert slik at du kan gjennomføre hver av dem uten å måtte sette deg inn i de etterfølgende.

Lykke til!

INNLEDNING

Koherens, en kvalitativt tilnærming

Matematikken er fysikkens språk, sies det. Vi har hittil i kurset beskrevet bølger med perfekte harmoniske matematiske funksjoner (eller i det minste periodiske bølger). Du kjenner antakelig også til at lys ofte blir beskrevet som fotoner med energi $E = hf$ (velkjente symboler) som om både energi og frekvens er veldefinerte størrelser (“skarpe størrelser”). Men alt dette er forenklinger. Det finnes ikke noe som helst i naturen som fullt ut matcher en matematisk harmonisk funksjon! Det er nå på tide å utvide vårt ordforråd og begrepsverden.

I denne prosjektoppgaven skal vi studere noen aspekter knyttet til mer realistiske bølger. Vi vil da betrakte fluktasjoner i bølgene over tid, og vi vil bruke en verktøykasse bestående av fouriertransformasjon, waveletanalyse og flere statistiske analysemetoder.

Fenomenene vi skal jobbe med gjelder for bølger generelt, enten det er lydbølger, radio-bølger, lys eller for den saks skyld også overflatebølger på vann. Vi skal bruke numeriske metoder i beregningene våre, og da er det enklest å jobbe med lydbølger. Vi kan da forholdsvis lett forholde oss til de bølgelengdene og periodetidene og geometrier vi kommer borti. Ved å betrakte lydbølger slipper vi også å ta hensyn til polarisering, noe som forenkler analysen. Vi legger opp til beskrivelser som er direkte testbare i praksis.

[♠ ⇒ Siden bølgefenomener knyttet f.eks. til lys og lyd har svært ulike tids- og lengdeskalaer, kunne det vært aktuelt å bruke såkalte “dimensjonsløse variable” for å behandle flere fenomener innen samme

formalisme. Vi har likevel bevisst valgt å ikke gjøre det. Brukes dimensjonsløse variable, må man forholde seg til mer abstrakte størrelser enn ved bruk av direkte fysiske mål. Det er allerede betydelige utfordringer med å få en god forståelse for de fenomenene vi skal ta opp, og da er det greit i første omgang å bruke en så enkel og konkret beskrivelse som mulig. Vi kommer imidlertid i siste del av oppgaven til å anvende den lærdommen vi gjør oss også på noen få fenomener knyttet til lys.

← ♠]

Generelle trekk for reelle bølger

Ulike bølgekilder

Tenk deg en lydkilde. Det kan være gjøk som synger sitt “ko-ko”, eller det kan være et fossefall som gir sitt vedvarende sus eller drønn. Lyden fra fossen sier vi er et “stasjonært signal” fordi lyden vi hører “nå” ikke oppleves som forskjellig fra det vi hørte for to sekunder siden eller for to timer siden. Gjøkens “ko-ko” har ikke samme karakteristikk, og kalles ikke-stasjonært. Det er ofte vanskelig å bedømme om et signal kan karakteriseres som stasjonært eller ikke, for det er en glidende overgang mellom disse. I vår oppgave skal vi betrakte stasjonære signaler, men du vil da nettopp se at det byr på utfordringer for våre forestillinger. I et vedlegg i slutten av dette skrivet er enkelte aktuelle ord og uttrykk forklart, f.eks. ordet “stasjonær”.

En kilde for bølger kan bestå av én enkeltkilde, flere enkeltkilder som er uavhengige av hverandre eller flere enkeltkilder som avhenger av hverandre. Frekvensen til kildene kan være vidt fordelt over et mer eller mindre bredt frekvensområde. Iblant er frekvensområdet så lite at vi kaller bølgen for “monokromatisk”, men alle reelle fysiske bølger har noe variasjon i frekvens etter som tiden går.

Et eksempel på en stasjonær bølge fra én kilde, er en enslig trompetist som spiller vedvarende på en tone. Lydbølgen kan bare sies å være stasjonær i den perioden trompetisten spiller (max et minutt eller så før hun må trekke pusten).

Et konstruert eksempel på et stasjonært signal med flere uavhengige kilder, men nokså smalt frekvensområde, er en rekke personer som spiller en vedvarende C på hver sin trompet. De tar korte pustepauser innimellom, men det skjer uavhengig av hverandre, slik at vi ikke har lett å oppdage hver enkelt pustepause dersom mange trompetister bidrar til sumsignalet. Selv om frekvensen til lyden alle produserer er i nærheten av hverandre, er det ingen sammenheng mellom fasene til lydene de genererer. Fasen er vilkårlig, eller varierer stokastisk, som vi iblant sier (se vedlegget).

En laser er et eksempel på en bølgekilde med mange enkeltkilder (atomer eller molekyler), men der vi har såkalt “stimulert emisjon” som gjør at enkeltatomene sender ut sitt lysbølgebidrag i fase med bølgene som har bygd seg opp i lasermaterialet i oppstartsperioden

for laseren. Laseren gir oss “koherent lys” sier vi, og vi skal se nærmere på hva dette betyr senere i denne oppgaven.

Lyset fra en stjerne er et eksempel på en helt annen prosess for lysutsending. Her vil plasmaet i noen deler av stjernen sende ut sitt lysbidrag totalt uavhengig av lysutsendelsen til plasmaet som ligger kilometervis unna det førstnevnte. Frekvensene til de ulike bidragene er til dels svært forskjellige, men de følger statistisk sett fordelingen som gjelder for såkalt “sort stråling” (stråling fra varme “legemer”). Vi tenker oss at lokale deler av plasmaet sender ut lys i mer eller mindre lange pulser av elektromagnetiske felt (“fotoner”). Siden det er så kolossalt mange partikler i plasmaet i en stjerne, vil lyset fra stjernen likevel bli sammenhengende i tid. Lyset fra en stjerne sier vi er ikke-koherent.

Det er faktisk lyset fra en stjerne som er utgangspunkt for prosjektoppgaven vår. To engelske fysikere, Robert Hanbury Brown og Richard Twiss, brukte i 1956 de statistiske egenskapene til lyset fra stjernen Sirius for å bestemme stjernens fysiske størrelse. Den metoden de utviklet er elegant og har visse likhetstrekk med tankegangen bak interferens.

Det hører med til historien at den såkalte Hanbury Brown og Twiss-effekten utløste opphetede debatter innen fysikkmiljøet, siden effekten er nokså lett å forklare ved hjelp av bølgebeskrivelsen til lys, men er vanskelig å forklare ved partikkelbeskrivelsen. Roy Glauber klarte på 1960-tallet å gi en kvantemekanisk forklaring basert på fotoner, og han fikk Nobelprisen i fysikk i 2005 for denne bragden. Men siste ord er neppe sagt i denne sammenheng. Vi vil f.eks. i denne prosjektoppgaven vise at vi skal forvente å finne samme effekt også for lydølger, og for lyd i luft er det ikke akkurat naturlig å snakke om fotoner (eller fononer)!

Sum av bølger

Vi har tidligere sett at lydølger forplanter seg i luft som “trykkølger”. Luftmolekylene blir påtrykt en kollektiv bevegelse (som kommer på toppen av de individuelle Brownske bevegelsene til molekylene). Det er her viktig å innse at i et lite luftvolum kan det ikke eksistere mer enn én verdi for det lokale lufttrykket til enhver tid. Det kan passere mange lydølger gjennom dette luftvolumet, gjerne samtidig. Likevel har vi med en reell summasjon å gjøre slik at det bare er én verdi på lufttrykket i små volumelementer til enhver tid. Vi kan gjerne beregne hvilken verdi dette skal være og bruke en matematisk modell som innebærer en sum av flere bidrag. Likevel: Luftmolekylene kjenner ikke til vår matematiske modell. I virkeligheten kan det ikke både være mange og få luftmolekyler samtidig innenfor ett og samme lille volum.

Overflatebølger på vann illustrerer dette poenget fint. I figur 1 er det vist et fotografi av slike bølger. Plukker vi ut et lite område på overflaten (lite i forhold til bølgelengdene som er involvert), ser vi at vannoverflaten i dette området bare har ett utslag (i forhold til en stille overflate) ved det tidspunktet bildet ble tatt. Vi kan gjerne beregne dette utslaget dersom vi vet hvilke ulike bølger som passerer området på den tiden vi betrakter. Men selv



Figur 1: *Overlatebølger på vann kan være sammensatt av mange bidrag. Imidlertid har vannoverflaten bare én høyde på ethvert sted til enhver tid.*

om vi i vår modellering bruker en sum av flere bidrag, har naturen selv bare ett utslag til enhver tid i det lille området vi betrakter. Vi kan ikke gjenkjenne bølgebidragene eller forstå hva som foregår uten at vi sammenligner utslagene også i naboområdene og følger utviklingen i tid i alle nærliggende små områder.

For en elektromagnetisk bølge (f.eks. lys) får vi nøyaktig samme type argument: I et område som er lite i forhold til bølgelengden, er det til enhver tid bare én verdi på det elektriske og magnetiske feltet.

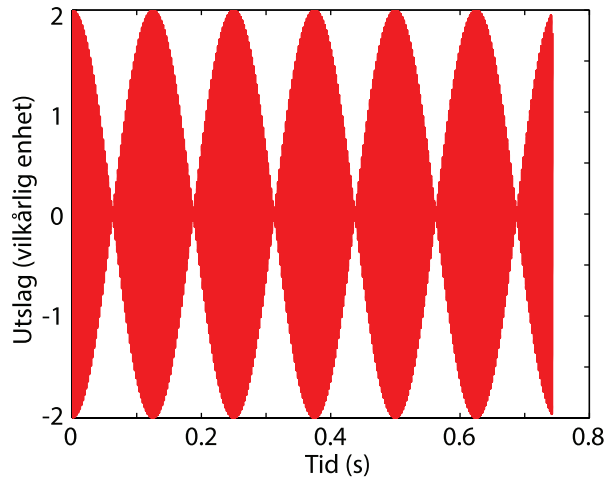
Det er denne tankegangen som ligger bak formalismen vi brukte for å beregne f.eks. interferens fra en dobbeltspalt eller et optisk gitter. Vi startet med veldefinerte elektriske felt i hvert lite volum (flate) i spalteåpningene, og vi beregnet summen av alle bidragene på ulike deler av skjermen.

Det høres kanskje uproblematisk ut at en sum av bølger er en sum av utslag i alle volumelementer til enhver tid. Men bildet kan fort bli mer problematisk. Allerede ved summering av to bølger får vi en utfordring. Vi kan beskrive dette slik:

$$y(t) = A \cos(\omega_1 t) + A \cos(\omega_2 t) \quad (1)$$

Matematisk sett har vi altså en enkel sum av to vedvarende harmoniske svingninger. Dette er et signal som i høy grad synes å være stasjonært. [Vi skal dog være forsiktig med å bruke ordet stasjonært om dette signalet, siden begrepet “stasjonært” egentlig hører hjemme i karakterisering av signaler fra stokastiske prosesser. Vi strekker strikken en del for å få fram et poeng som kommer til syne også i signaler fra stokastiske prosesser.]

La oss nå betrakte dette enkle signalet med ulike analysemetoder:



Figur 2: Tidspildet av en funksjon som er en sum av to sinuser med litt forskjellig frekvens, men samme amplitude (i tråd med ligning (1)).

a) Tidspildet av funksjonen

Figur 2 viser tidspildet til signalet i ligning (1) for $\omega_1 = 400$ Hz og $\omega_2 = 408$ Hz. Vi får en typisk beat-effekt (svevelyd-effekt), som også kommer fram ved den velkjente omskrivingen av ligning (1):

$$y(t) = 2A \cos\left(\frac{\omega_1 + \omega_2}{2} \cdot t\right) \cos\left(\frac{\omega_1 - \omega_2}{2} \cdot t\right)$$

b) Fouriertransformasjon

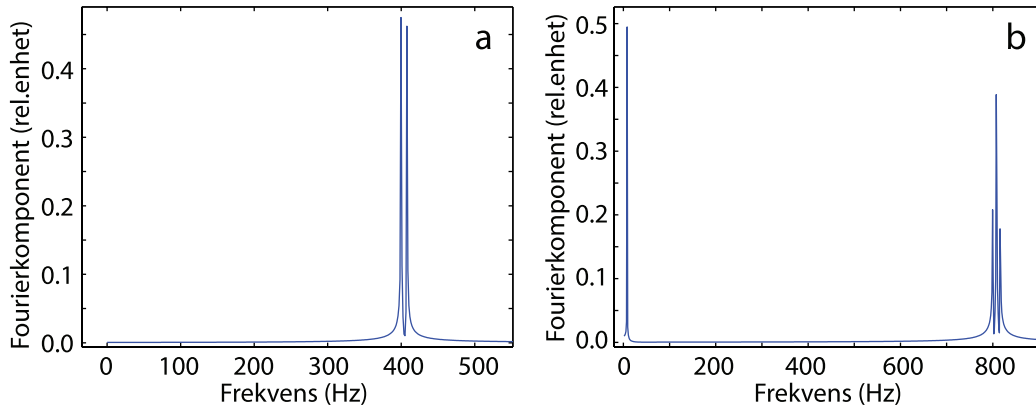
Venstre del av figur 3 viser den fouriertransformerte av y . Merk at dette resultatet er fullt ut i tråd med beskrivelsen i ligning (1). Hvor blir det av beat-frekvensen? Det finnes overhodet ikke noe antydning til denne i fourierspekteret til signalet.

Vi kan imidlertid gjøre et triks som får fram beatfrekvensen også i en fouriertransformasjon. Beregner vi først kvadratet av y i hvert punkt, og så tar en fouriertransformasjon, får vi resultatet i figur 2 høyre del. Her kommer beatfrekvensen fram i frekvensområdet nedenfor enkeltfrekvensene, og sumfrekvensen i området ovenfor enkeltfrekvensene. Det er også såkalte "sidebånd" til sumfrekvensen.

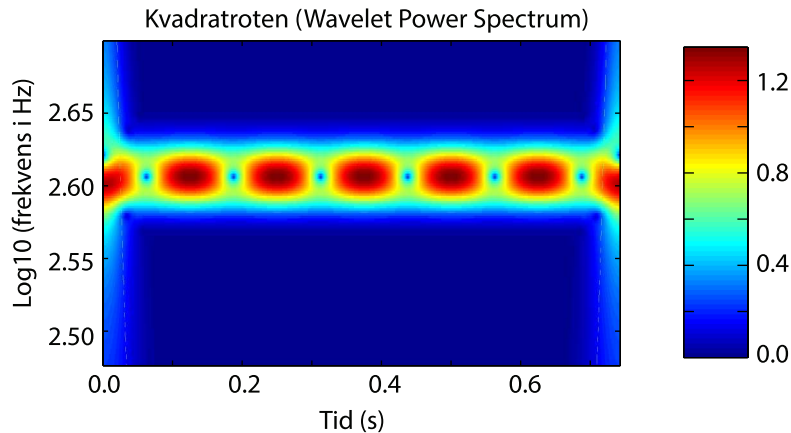
c) Wavelettransformasjon

Figur 4 viser waveletanalyse av signalet (omtrent optimalisert for dette konkrete signalet). Vi får her fram et bilde som stemmer godt overens med vår egen lydopplevelse. Men hadde vi valgt en ekstrem høy K-verdi i waveletanalysen, ville vi fått et resultat som hadde nærmet seg den enkle fourieranalysen.

Kan vi si at dette signalet er stasjonært? Det er jo stor forskjell på bølgen ved tiden 0.50 s sammenlignet med bølgen ved tiden 0.56 s. Vi unnlater å ta ordentlig stilling til dilemmaet



Figur 3: *Fourieranalyse av sumsignalet i ligning 1 i venstre del. Det er topper ved ω_1 og ω_2 . I høyre del er kvadratet (punktvis) av sumsignalet i ligning 1 analysert. Toppene befinner seg nå ved frekvensene $\omega_D = \omega_2 - \omega_1 = 8$ Hz, og ved $\omega_S = \omega_1 + \omega_2 = 808$ Hz og ved $\omega_S \pm \omega_D = 800$ Hz ($= 2\omega_1$) og 816 Hz ($= 2\omega_2$). Merk at beatfrekvensen $\omega_D \ll \omega_1$ og ω_2 .*



Figur 4: *Waveletanalyse med Morlet wavelets av tidssignalet fra ligning (1). ($K = 32$.)*

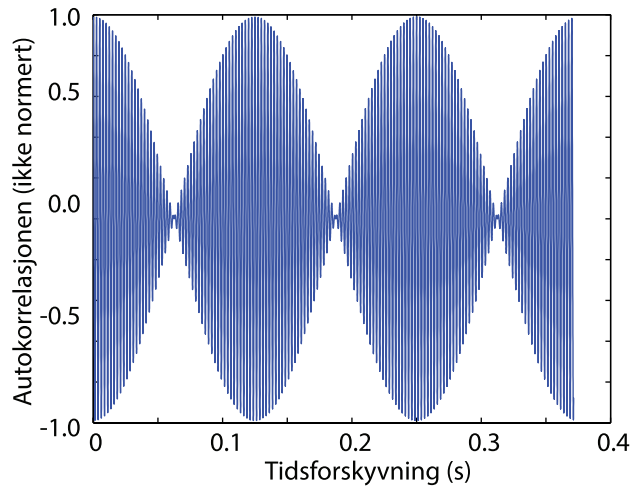
vi står overfor ved å si at bølgen er stasjonær for tider som er mye større enn halve periodetiden for beatsignalet $T_b/2 = 62.5$ ms, men ikke-stasjonær for tider mye mindre enn denne periodetiden.

d) Autokorrelasjonsfunksjonen

Det finnes en statistisk analysemetode som iblant er meget nyttig når vi skal utforske reelle bølger. Metoden er lite nyttig for enkle og ikke-stokastiske signaler slik vi har i ligning (1), men vi presenterer den allerede her siden du vil ha bruk for den i resten av oppgaven.

Autokorrelasjonsfunksjonen til et signal kan defineres ved det matematiske uttrykket:

$$R(\tau) = \int f(t) * f(t + \tau) dt / \int f^2(t) dt$$



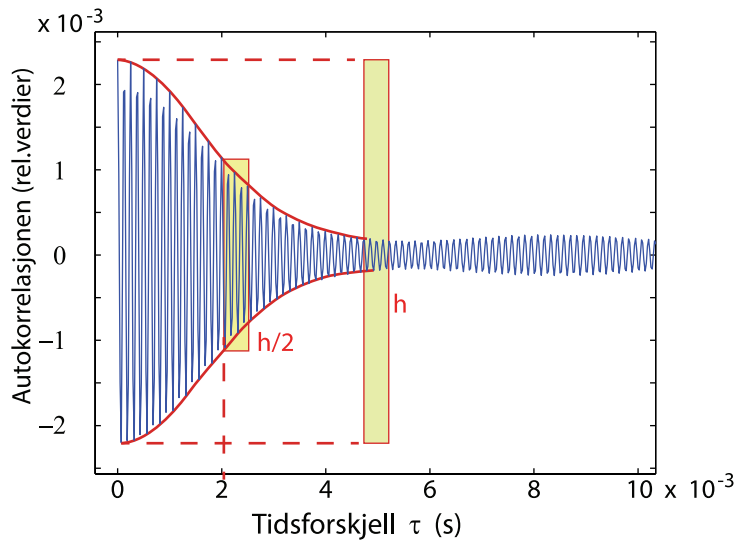
Figur 5: Autokorrelasjonsfunksjonen av en sum av to sinuser gitt i ligning (1). Merk at tidsstrengen nå bare er halvparten av hva den var i figur 2.

Autokorrelasjonsfunksjonen sjekker rett og slett hvor stor overensstemmelse det er mellom en tidssekvens av et signal og en tidssekvens som er litt forskjøvet i forhold til den første. Det sier seg selv at når $\tau = 0$, er det perfekt korrelasjon, og da vil $f(\tau) = 1$. Dersom en tidsforskyvning τ_2 er slik at $f(t) = -f(t + \tau)$, vil $R(\tau_2) = -1$. For en ren matematisk sinusfunksjon vil derfor autokorrelasjonsfunksjonen rett og slett bli en cosinus-funksjon.

For vårt signal gitt i ligning (1) ser autokorrelasjonsfunksjonen ut som vist i figur 5. Vi får også her et repeterende signal som ikke er særlig overraskende siden vi har en sum av to matematisk perfekte harmoniske signaler.

Det er vanlig å si at autokorrelasjonsfunksjonen angir hvor langt fram i tid vi kan forutsi hvordan signalet oppfører seg på en forutsigbar måte. Så lenge $R(\tau)$ viser verdier i nærheten av 1 i tallverdi, betyr det at det er stor grad av sammenheng mellom signalet “nå” med det signalet blir om τ sekunder.

Figur 6 viser autokorrelasjonsfunksjonen for et annet signal, - et signal som viser stor grad av variasjon etter som tiden går. Etter bare ca 5 millisekunder er autokorrelasjonsfunksjonen kommet ned til et “støynivå” som er “nær null”. For dette signalet er det en avtakende korrelasjon (i tallverdi) mellom signalet “nå” og det som kommer noen få millisekund senere. I denne perioden blir R iblant negativ - noe vi skal forvente siden det tross alt er en bølge vi analyserer (selv om den er uregelmessig) .



Figur 6: Eksempel på en autokorrelasjonsfunksjon av en stokastisk bølge. Bare de få første millisekundene av funksjonen er tatt med (resten viser bare tilfeldige variasjoner). Tiden det tar fra autokorrelasjonen har sunket fra sin første verdi til halvparten av denne verdien, vil vi kalle koherenstiden. I dette tilfellet er den omtrent 2 ms.

DEL 1: TILFELDIGE BØLGER

Det konkrete arbeidet i prosjektoppgaven består av flere tilsynelatende forskjellige temaer (Del 1-5). I Del 1 skal du starte med å lage en verktøykasse av programmer som går igjen i senere deler. Du skal også studere karakteristiske sammenhenger mellom bredder på frekvensfordelingen til et signal og koherens. Men la oss altså starte med å programmere funksjoner vi har bruk for i resten av oppgaven.

I. Generering av funksjoner med tilfeldig tidsutvikling (random functions)

Suset fra en foss er et eksempel på et lydsignal som varierer tilfeldig over tid. Lyset fra en stjerne er likeså. Men hvordan skal vi modellere et slikt signal numerisk?

For stjernen kan vi tenke oss at vi kan modellere lyset som en sum av svært mange lyssignaler som kommer fra ulike atomer (plasma) ved tilfeldige tidspunkt. Hvert enkelt lyssignal kunne modelleres som bølgepakker, hver med sin senterfrekvens og bredde på omhyllingskurven. Vi kunne eksperimentere med ulike modeller og sett hva som svarte best med det vi kan observere.

En slik form for modellering tar mye regnemaskintid, og i praksis fører den i vår sammenheng ikke til merkbare forskjeller sammenlignet med fra en langt mer effektiv måte å lage et

signal med vilkårlig tidsutvikling (egen uttesting april 2014). Den effektive modelleringen kan vi gjøre ved å generere et random frekvensspekter og så bruke invers fouriertransformasjon for å få et tidsbilde som har de karakteristiske trekkene vi ønsker.

Prosedyren er da som følger:

Vi så i figur 3 at den fouriertransformerte av summen av to harmoniske funksjoner er to linjer i frekvensspekteret. I figuren anga vi bare absoluttverdien av den fouriertransformerte. Vi vet at vi kan få tilbake det opprinnelige signalet ved en invers fouriertransformasjon. I så fall må vi ta vare på faseinformasjonen - det vil si realdel og imaginærdel hver for seg for begge frekvensene.

Det betyr at vi kan STARTE med et frekvensspekter for så å generere et signal i tidsbildet med de frekvenskarakteristikkene vi ønsker oss. Vi kan f.eks. lage et signal der alle mulige frekvenser i et visst frekvensområde inngår. Vi kan velge hvor sterkt bidrag hver enkelt frekvens skal ha, og hvilken fase hver av frekvensene skal ha.

Når vi skal modellere dette numerisk, må vi velge frekvensområde og samplingsfrekvens. Vi skal beskrive lyd, og et naturlig valg er da standard samplingsfrekvens for CDer, nemlig 44100 Hz. [I siste del av oppgaven kommer enkelte detaljer bedre fram om vi velger en høyere samplingsfrekvens. Vi kunne da med fordel brukt 88200 Hz, som forøvrig blir brukt også i enkelte profesjonelle lydopptak i dag.]

Vi vet fra før at oppløsningen på frekvensaksen avhenger av hvor lang tid samplingen strekker seg over. Vi vil velge relativt lange datastrenger for å få en god frekvensoppløsning, f.eks. $2^{16} = 65536$ punkter.

Spesifikasjon av hvordan programmet skal fungere:

- Som allerede nevnt: Signal som genereres skal bestå av 2^{16} punkter, og samplingsfrekvensen skal være 44100 (i alle fall i starten).
- Vi må velge frekvenskarakteristikken. For lys fra en stjerne ville vi valgt sort-stråling-karakteristikken for en egnet fargetemperatur. Vi skal modellere lydsignaler, og velger av praktiske grunner en gaussisk amplitudfordeling med en gitt senterfrekvens og en gitt bredde på fordelingen (hvor amplituden på signalet er sunket til $1/e$ av hva den er ved senterfrekvensen, statistisk sett). For å få vilkårlighet inn i bildet, velger vi for hver frekvenskomponent en vilkårlig amplitude mellom null og verdien gitt av den gaussiske fordelingen. (Én måte å angi en gaussisk fordeling er vist i vedlegget, og et eksempel på en Matlab-kode er gitt på en av de siste sidene i dette skrevet.)
- For å igjen få mest mulig vilkårlighet inn i signalet, må vi velge en vilkårlig fase ϕ mellom 0 og 2π for hver frekvenskomponent. Realdelen av frekvenskomponenten blir da amplitudeverdien multiplisert med $\cos \phi$ og imaginærdelen blir samme amplitudeverdien multiplisert med $\sin \phi$.

- For at et fourierspekter skal ende opp som en reell funksjon etter en invers fouriertransformasjon, må komponentene følge kriteriene for folding:
Dersom FT-spekteret består av $2^n = N$ punkter y_k hvor $k = 1, 2, 3, \dots, N$, representerer første punkt y_1 middelveien av tidssignalet etter en invers FT, og kan settes lik null. Punktene $y_2 - y_{(N/2)+1}$ er de komplekse frekvenskomponentene mellom $f_1 = 1/T$ til $f_{(N/2)+1} = F_s/2$ hvor F_s er samplingsfrekvensen. På grunn av speilingen, må punktene $f_{(N/2)+1}$ til f_N være kompleks konjugerte til frekvenskomponentene nedenfor halve samplingsfrekvensen (se figur 15 og 16 i 2014-versjonen av kapittel 4), men altså speilet om halve samplingsfrekvensen. Det betyr at $f_{(N/2)+1}$ selv må være reell.
- Det er viktig å sjekke at signalet etter en invers fouriertransformasjon blir reell (at imaginærkomponentene er neglisjerbare i forhold til realkomponentene). Dersom det ikke er tilfelle, har man ikke tilfredsstilt kravene som ligger i speilingen. Når alt er som det skal, bør man velge å bare bruke realkomponenten videre for å slippe å gjøre dobbel jobb til ingen nytte.
- Plott også frekvensfordelingen for å sjekke at amplitudfordelingen er gaussisk (omhyllingen til de vilkårlige amplitudene), og at senterfrekvensen og bredden på fordelingen er slik den skal være. Sett som alltid på riktige tall og tekst langs aksene.

Du bør lage en frittstående funksjon (egen fil) som genererer det vilkårlige signalet, og et separat hovedprogram hvor du velger parametre, overfører disse ved kall til funksjonen, og henter tilbake det vilkårlige signalet (for senere analyse). I vedlegget er det gitt et eksempel på et hovedprogram.

Som vanlig anbefaler vi at du IKKE forsøker å skrive hele programmet i en jafs. Start med det helt enkle og test ut at ting fungerer som du ønsker underveis før du legger til enda flere detaljer. Dette gjelder såvel når du skriver en funksjon og når du bygger opp hovedprogrammet.

II. Program for waveletanalyse

Anvend programmet gitt i kapittel 13 eller skriv ditt eget program for waveletanalyse med bruk av Morlet wavelets etter de retningslinjene som er gitt i kapittel 13. Test programmet f.eks. med fuglelyd-filen tilgjengelig fra kurswebsidene (sammenlign resultatet med tilsvarende figur i kapittel 13). Vær sikker på at du behersker å sette parametrene slik du ønsker de skal være!

Detaljer som kan skjære seg ved bruk av funksjonen `WL1.m` gitt side 471 og 472 i kapittel 13 (2014-versjonen):

- Dersom du ved bruk av funksjonen får feilmeldingen “Matrix dimentions must agree”, kan det hende du bør skifte fra kollonne-vektor til rad-vektor på input-vektoren h . Prøv “ $h = \text{transpose}(h)$ ” for dette skiftet enten i det kallende programmet eller like etter start av WL1.m.
- Dersom det oppstår en feil med referanse til koden mellom “% Foreta en markering i ...” til “% Plotter waveltdiagrammet”, kan du rett og slett kommentere ut disse 8 linjene med kode. Waveletdiagrammet vil i så fall mangle markeringslinjen for randproblem langs høyre og venstre kant, men det kan ofte aksepteres. Ta inn disse linjene i programmet igjen når du analyserer signaler hvor problemet ikke oppstår.
- Dersom det oppstår en feil med referanse til koden mellom “% Reduser filstørrelse ved å ... ” og “% Foreta en markering i ...” kan du forsøke å endre tallet 24 i linjen hvor P berengnes, til halvparten eller det dobbelte, eller rett og slett sette P lik 1 (ingen reduksjon i filstørrelse for plottene).

Programmet for waveletanalyse bør også være en separat funksjon du kan aktivere fra et hovedprogram.

III. Program for beregning av autokorrelasjonsfunksjonen

En første ordens korrelasjonsfunksjon av et signal $f(t)$ kan defineres slik (mangler da normalisering, hvilket betyr at $F(0) \neq 1$):

$$F(\tau) = 1/T \int_0^T f(t)f(t + \tau)dt \quad (2)$$

Er signalet beskrevet ved diskrete tidspunkt, ser uttrykket ut som følger:

$$F(\tau_j) = 1/M \sum_{i=1}^M f(i)f(i + j) \quad (3)$$

hvor tiden $\tau_j = j/F_s$ der j er et heltall og F_s er samplingsfrekvensen.

Vi multipliserer med andre ord bølgen med en tidsforskjøvet bit av samme bølge, punkt for punkt, og summerer over alle punktene og dividerer med antall punkter i beskrivelsen.

Merk at når $i = M$ i ligning 3, vil $(i + j) > M$. Det betyr at M ikke kan være like stor som størrelsen N på arrayen som beskriver signalet f . Dersom du lar $M = N/2$ vil du kunne beregne $F(\tau_j)$ for τ_j verdier fra null til halve tiden som f er beskrevet over.

Siden korrelasjonsfunksjonen vi har definert innebærer bare ett signal, hvor vi sammenholder en del av signalet med en annen del av det samme signalet, kalles den for “autokorrelasjonsfunksjonen”.

Lag en separat funksjon som kan beregne og plote autokorrelasjonsfunksjonen til et signal (som overføres som parameter til funksjonen fra hovedprogrammet). Anvend ligning (3). Test om programmet fungerer som det skal ved å bruke testfilen lagt ut på kurswebsidene. Resultatet bør bli identisk med figur 6.

Husk at autokorrelasjonsfunksjonen ofte faller ned til et “støynivå” etter svært kort tid. Det er derfor vanligvis nødvendig å zoome inn betydelig den aller første delen av kurven for å få fram det som er interessant.

Koherenstid

Maksimumsverdien for autokorrelasjonsfunksjonen er vanligvis første punkt $F(\tau)_{max} = F(0)$. $|F(\tau)|$ avtar (gjørne på en oscillerende måte) etter som τ øker, men vil oftest nå et “støynivå” etter svært kort tid. Omtrent den tiden det tar for $|F(\tau)|$ å ha sunket til ca halvparten av startverdien $F(0)$, kaller vi signalets koherenstid (tidmessig koherens, engelsk: temporal coherence).

Iblant er støynivået i autokorrelasjonsfunksjonen omtrent like stort som den første, interessante delen. Det skyldes at tidsstrengen vi analyserer er så kort i forhold til signalets karakter at tilfeldigheter ikke dempes godt nok. Den første delen av autokorrelasjonsfunksjonen for bølger vil alltid ha en struktur av typen vist i figur 6, mens variasjonene etter dette vil variere etter som vi velger ulike utsnitt fra en og samme tidsstreng. Vi konsentrerer oss derfor bare om den innledende delen der omhyllingskurven omtrent har en eksponentielt avtakende form.

Studier av tilfeldige signaler; koherens vs bredde i frekvensfordeling

Du bør nå ha tre ferdig uttestede funksjoner tilgjengelig som

1. kan generere et tilfeldig signal gitt en del parametre,
2. foreta en waveletanalyse av signalet, og
3. beregne autokorrelasjonsfunksjonen til signalet.

Du må også delvis ha på plass et hovedprogram hvor du angir parameterverdier og kaller funksjonene etter behov.

Et eller annet sted må det også plottes signalet du jobber med i tidsbildet (dvs $f(t)$), og du må kunne plote frekvensspekteret av signalet, det vil si den fouriertransformerte av signalet. Disse plottene kan f.eks. gjøres i funksjonen som lager det tilfeldige signalet, eller plotting kan gjøres i hovedprogrammet.

Vi er nå spent på hva analysene kan gi for ulike tilfeldige signaler. Vi skal nå studere hvordan koherenslengden/koherenstiden avhenger av senterfrekvens og bredde i frekvensfordelingen

for ulike tilfeldige signaler.

Kjør alle fire analysene for tilfeldige signaler generert med følgende parametre:

Forslag til parametre:

a) Vi skal først teste hvordan koherenstiden avhenger av senterfrekvensen til signalet, men lar bredden i frekvensfordelingen være uforandret. Parametre som da egner seg er: Fullverdibredde i frekvensfordelingen: 50 Hz (fra nedre frekvens som har amplitudeverdi $1/e$ av max til øvre frekvens som har samme amplitudeverdi, statistisk sett). Gaussisk frekvensfordeling. Senterfrekvens etter tur: 100, 1000 og 10000 Hz. Bruk fullt sett analyser (tidsbildet, frekvensbildet, waveletanalyse og autokorrelasjonsfunksjonen). Den omtrentlige koherenstiden bestemmes for hvert av disse tilfellene ut fra prosedyren gitt i figur 6. Det kan forøvrig være et triks å plote absoluttverdien av autokorrelasjonsfunksjonen for å lette avlesningen av ved hvilken tid autokorrelasjonsfunksjonen (konvolutteringen) har sunket til halvparten av den initielle verdien. K-verdien for waveletanalysen kan være om lag 48 for alle disse kjøringene. Husk å justere frekvensområdet for waveletanalysen etter som du endrer på frekvensen til signalet!

b) Vi skal dernest teste hvordan koherenstiden avhenger av bredden av frekvensfordelingen, mens senterfrekvensen holdes uforandret. Parametre som da egner seg er: Senterfrekvens 10000 Hz (10 kHz). Fullverdibredde i frekvensfordelingen etter tur: 0.5, 10, 100, 1000 og 5000 Hz. Også her skal alle analyser gjennomføres og man velger ut eksempler på plot som man vil ha med i rapporten. Koherenstiden bestemmes for hvert av tilfellene. Det kan med fordel lages et plot av koherenstiden vs bredden av frekvensfordelingen. K-verdien for waveletanalysen kan være om lag 48 for alle disse kjøringene, men du kan også gjerne justere verdien etter hva du mener er gunstig.

Resultater/vurderinger som bør med i rapporten:

For disse kjøringene ber vi deg:

- Plot et lite utsnitt av tidssignalene du genererer. Ta med to til fire eksempler på tidsSignaler i rapporten. Forsøk å beskrive likheter og forskjeller mellom de ulike signalene som er generert alt etter som parametrene endres.
- Ta med minst et par eksempler på frekvensspekteret til de signalene du genererer.
- Ta med minst tre eksempler på waveletanalyser i rapporten, gjerne med passe utsnitt av totalen. Beskriv hva du ser. I prinsippet har du laget signalene ved å addere mange signaler med ulik amplitude, men med nøyaktig samme amplitude for hele tidsrommet signalet varer. Det du observerer avviker tilsynelatende fra et slikt bilde. Hvordan kan det ha seg? Husk å justere frekvensgrensene for waveletanalysen slik at du får med deg det som er interessant (og ikke svært mye mer).

- Bestem omtrentlig koherenstiden, dvs hvor lenge autokorrelasjonsfunksjonen er tydelig forskjellig fra bakgrunnsnivået vi finner ved store τ . Ta med minst et par eksempler i rapporten på plot av autokorrelasjonsfunksjonen for de signalene du studerer. Husk å zoome inn på små τ for å finne den interessante delen av funksjonen! Ellers er det vesentlig å forsøke å finne en lovmessighet mellom parametre brukt for generering av det tilfeldige signalet og koherenstiden for signalet! Husk at vi jobber med statistiske signaler og får da på langt nær så presise verdier som når vi jobber med analytisk matematikk alene. Hvor vil du sette skillet mellom “koherent signal” og “ikke-koherent signal”? Drøft problemstillingen.

DEL 2: HBT EFFEKTEN - EN FORSIKTIG START

I del 1 så du forhåpentligvis at når vi summerer mange uavhengige bølger spredt over et visst frekvensbånd, får vi en tilfeldig bølge. I denne tilfeldige bølgen er de ulike frekvensene i matematisk forstand (fourieranalyse) hele tiden konstant til stede, men summasjonen fører til en stadig veksling i tid mellom hvilke frekvensbidrag som nuller hverandre ut og hvilke som adderes konstruktivt. Resultatet kommer flott ut i et waveletdiagram, og vi kan bli lurt til å tro at vi har med “partikler” (fotoner) å gjøre, mens det overhodet ikke finnes noen partikler i vår beskrivelse, bare en rett fram summasjon av bølger med en spredning i amplitude, frekvens og fase.

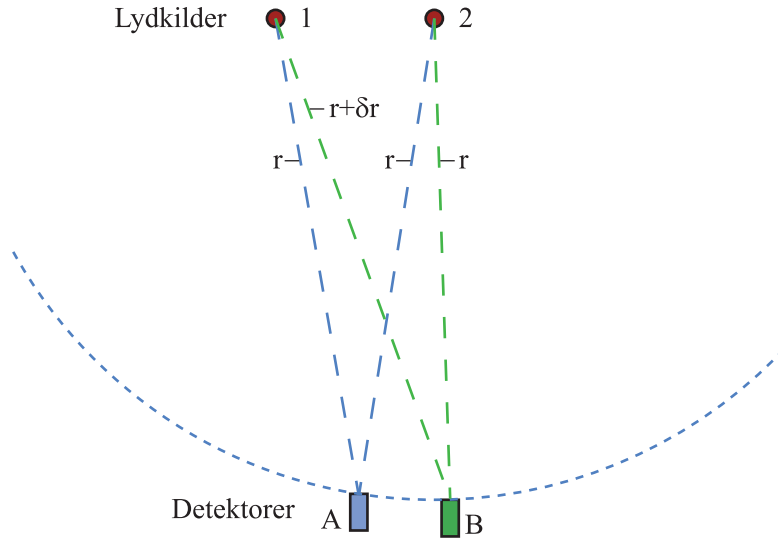
Den tilfeldige variasjonen med tiden kan vi benytte til elegante målemetoder. Hanbury Brown og Twiss gjorde dette for første gang i 1954 (radioastronomi) og 1956 (synlig lys astronomi). I dag brukes relaterte, nydelige teknikker i studier av statistiske detaljer i vidt forskjellige systemer, blant annet for å følge hastigheten til sammenklumping av blodplater (aktuelt for å forutsi fare for blodpropp). Dette er nok et eksempel på at tilsynelatende unyttig grunnforskning innen fysikk kan få praktisk anvendelse.

For å forstå effekten som ligger bak HBT-effekten skal vi starte ganske forsiktig. Effekten framkommer når vi summerer to (eller flere) tilfeldige bølger, og iblant forsinker en av bølgene i forhold til den andre. Vi sammenligner så bølgen for det tilfellet at de to kildene adderes uten tidsforsinkelse med bølgen med tidsforsinkelse.

Figur 7 viser en prinsippskisse for hva vi skal modellere. Vi har to lydkilder, 1 og 2, med ulike vilkårlige signaler $k_1(t)$ og $k_2(t)$. De to kildene har samme avstand r til en av to detektorer (detektor A), og signalene fra 1 og 2 kan derfor summeres uten en relativ tidsforsinkelse. Signalet inn på detektor A kaller vi $f(t)$, og vi har:

$$f(t) = k_1(t + \Delta t) + k_2(t + \Delta t)$$

hvor Δt er tiden bølgen bruker på å gå avstanden r .



Figur 7: Geometrien som ligger bak HBT-effekten. Se teksten for detaljer.

En annen detektor (B) har samme avstand r til kilde 2, men variabel avstand til kilde 1. Avstanden til kilde 1 er $r + \delta r$. Total bølge som når detektor B er derfor en sum av bølgen fra kilde 2 og en (relativt sett) tidsforsinket bølge fra kilde 1. Signalet inn på detektor B kaller vi $g(t)$, og vi har:

$$g_{\delta t}(t) = k1(t + \Delta t + \delta t) + k2(t + \Delta t)$$

hvor δt er tiden bølgen bruker på å gå avstanden δr . Vi har satt en indeks på g slik at den angis som $g_{\delta t}(t)$ for å markere at g vil avhenge av hvor detektor B plasseres langs den stiplede sirkelbuen på figur 7.

Vi opererer med tilfeldig valg av nullpunkt for tid, så det er egentlig ikke noe vits i å trekke med seg Δt i uttrykkene våre. Vi velger derfor å skrive uttrykkene for f og g slik:

$$f(t) = k1(t) + k2(t)$$

$$g_{\delta t}(t) = k1(t + \delta t) + k2(t)$$

Når funksjonene $k1$ og $k2$ er beskrevet i diskrete tidspunkt, vil denne siste ligningen måtte skrives slik:

$$g_{\delta n}(i) = k1(i + \delta n) + k2(i)$$

hvor δn nå svarer til tidsforskyvningen δt på følgende måte: $\delta n = \text{round}(F_s \cdot \delta t)$. Vi får altså diskrete hopp i tid på grunn av at samplingsfrekvensen F_s er endelig.

Signalene f inn på detektor A skal så sammenlignes med signal g inn på detektor B. Sammenligningen bør foretas på flere måter:

1. Beregn den såkalte krysskorrelasjonen mellom signalene som funksjon av den relative tidsforsinkelsen. Tidsforsinkelsen kan beregnes ut fra δr og bølgehastigheten, men i Del 2 hopper vi over disse detaljene og bare vurderer krysskorrelasjonen som funksjon av *hvor mange punkter* δn vi forskyver utsnittet vi benytter fra signalet $k1$ fra kilde 1 når sumsignalet $g_{\delta n}(i)$ på detektor B skal beregnes.

Vi vil definere kryss-korrelasjonsfunksjonen på følgende måte:

$$CC(\delta n) = \frac{1}{N} \sum_{m=1}^N |f(m)g_{\delta n}(m)|. \quad (4)$$

hvor f og g er definert ovenfor. Vi har av praktiske grunner valgt å bruke absoluttverdien før summasjon (det er ikke nødvendigvis et beste valg).

Beregn og plot hvordan krysskorrelasjonen endrer seg når δn variere fra 0 til 2000.

2. Studer waveletdiagrammene av f og $g_{\delta n}$ for minst tre ulike valg av δn (velg verdier ut fra plottet du fikk i forrige punkt). Ta med minst to av disse i rapporten. Studér waveletdiagrammene nøye (evt utsnitt) slik at du er sikker på at du skjønner hvorfor krysskorrelasjonen endrer seg med δn . (Bruk $K = 48$ eller en bedre verdi dersom du ønsker å perfeksjonere, men ikke bruk tid på finjustering!!!)

Foreshår følgende parametre:

- a) Generer to ulike tilfeldige signaler svarende til $k1$ og k på samme måte som i DEL 1, hvert signal med senterfrekvens 10000 Hz og fullverdibredde 1000 Hz.
- b) La "signal inn på A" (dvs f) være summen av de første 60 000 punktene av $k1$ pluss de første 60 000 punktene av $k2$.
- c) La "signal inn på B" (dvs g) være summen av de første 60 000 punktene av $k1$ pluss punktene $1 + \delta n$ til $60000 + \delta n$ fra $k2$. (Da kan δn ikke overskride vel 5000, men det er heller ikke aktuelt.)

Analog analyse, men nå av intensiteter: Square law detector

Vi skal om litt gjenta akkurat de samme to punktene som ovenfor, men nå ved å bruke intensiteter til signalene f og $g_{\delta n}$ i stedet for signaler direkte. Begrunnelsen for dette er følgende:

Vi har betraktet tidsvariasjonen til signalene f og g inn på detektor A og B både ved hjelp av waveletanalyse og korrelasjonsbestemmelser. I analysen har vi behandlet alle signaler lineært, det vil si at signalverdien svinger mellom positive og negative verdeier.

I vår sammenheng er det helt ok, for en mikrofon ville kunne gi en tidsvariasjon av våre lydsignaler som var positiv og negativ i forhold til ikke-noe-signal. Tidsvariasjonene som er aktuelle ved lyd i luft er så langsomme at vi kunne føre (forsterkede) mikrofonsignaler over en elektrisk kabel over flere hundre meter før analyse uten å få problemer med stabilitet osv.

I Hanbury Brown og Twiss eksperimentet fra 1956 betraktet man lys fra stjernen Sirius. Ingen detektorer har hittil klart å følge den høye frekvensen som lysbølger har. Det er umulig å få fram signaler fra lys på akkurat en slik måte vi har jobbet med dem til nå.

Detektorer for lys fungerer som square law detectors" (utsignalet er proporsjonalt med intensiteten på bølgen som kommer inn). Men utsignalet klarer slett ikke å endre seg så raskt som periodetiden til inputsignalet. Detektorene foretar derfor i praksis en form for filtrering av intensitets-signalet, og bare forandringer i tid som skjer mye saktere enn den virkelige bølgen finnes i signalet kommer ut fra detektorene. I HBT-eksperimentet var båndbredden på detektorer og forsterkere 5 - 45 MHz.

Signalene i HBT ble forsterket opp og ført sammen for beregning av krysskorrelasjonen. Når frekvensen etter filtrering var så lav som den var, kunne signalene overføres over flere hundre meter uten at det gikk ut over stabiliteten (i alle fall for signalene i radioastronomi-arbeidet). Det betyr at detektor A og detektor B i HBT kunne ha en innbyrdes avstand på flere hundre meter, og likevel kunne man foreta analyse av kryss-korrelasjonen mellom detektorsignalene.

Selv om vi strengt tatt ikke har bruk for å bruke intensitets-korrelasjoner for vårt arbeid med lydbølger, vil vi likevel legge inn denne detaljen i våre numeriske beregninger. Hensikten er bare å vise hvordan dette faktisk fungerer. Vi trenger da å lage et filter som bare slipper gjennom langsomme tidsvariasjoner: et såkalt lavpassfilter.

Lavpassfilter kan vi lage slik: Vi fouriertransformerer signalet, fjerner frekvenskomponenter i frekvensbildet vi ikke vil ha, og så gå tilbake til tidsbildet igjen vha en omvendt fouriertransformasjon.

Lag din egen lavpassfilterfunksjon på følgende måte: Funksjonen får overført et signal $u(n)$ fra det kallende programmet. Arrayen u er definert i N punkter, og u , N og samplingsfrekvensen F_s må være inputparametre, såvel som øvre frekvens f_{max} som skal beholdes uforandret.

I programmet fouriertransformerer du u . Du beregner så hvilken posisjon i fourierspekteret grensefrekvensen f_{max} svarer til, og så nuller du ut et område av frekvensspekteret sentrert rundt punktet $(N/2 + 1)$. Da vil frekvensene under f_{max} (og de speilede frekvensene øverst i spekteret) blir bevart. Etter en invers fouriertransformasjon kan det ferdig filtrerte signalet sendes tilbake til det kallende programmet. Bruk bare realdelen av signalet etter den inverse fouriertransformasjonen (imaginærdelen skal allikevel være null dersom du har gjort ting rett).

Husk å IKKE bruke abs-verdier i frekvensspekteret underveis. Realdel og imaginærdel av frekvensspekteret (etter den innledende fouriertransformasjonen) må være uforandret for de frekvenskomponentene du beholder. Sjekk at funksjonen fungerer som den skal før du bruker den i vår prosjektoppgave-sammenheng.

Nå er det på tide å faktisk gjennomføre analysen på intensitetsvariasjoner i tid.

Intensitetsvariasjon i tid får vi ved å ta punktvis kvadrering av signalene inn til detektorene:

$$f\text{-signal-intensitet} = f.*f$$

og tilsvarende for g.

Utsignalet fra detektor A vil da være

$$A2LP = \text{lavpassfilter}(f.*f)$$

og tilsvarende for g

Vi skal nå sammenligne $A2LP = \text{lavpass}(f.*f)$ med $B2LP_{\delta n} = \text{lavpass}(g_{\delta n}.*g_{\delta n})$, slik:

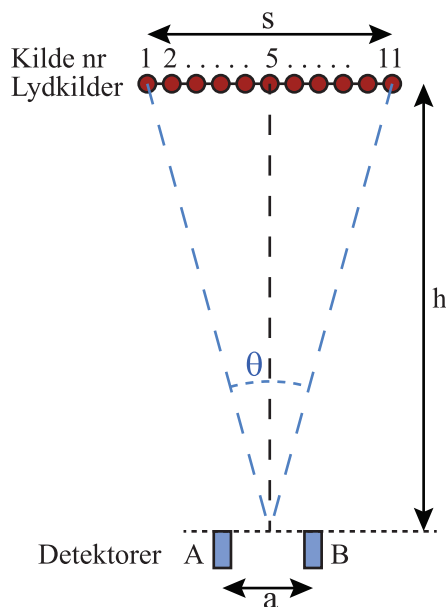
1. Beregn og plot hvordan den såkalte krysskorrelasjonen for $A2LP$ og $B2LP_{\delta n}$ varierer med δn (der som før δn svarer til den relative tidsforskyvningen i signal 1 til detektor B). La igjen δn variere fra 0 til 2000.
2. Analyserer også signalene $A2LP$ og $B2LP_{\delta n}$ med waveletanalyse for ca tre ulike δn -verdier (valgt ut fra plottet i forrige punkt). Bruk $K = 48$ eller et bevist valg hos deg selv. Husk å justere hvilke frekvenser du vil bruke i waveletdiagrammet nå etter at det er brukt et lavpassfilter. Vi vil bruke et filter som bare slipper gjennom vesentlig lavere frekvens enn det frekvensområdet som ble brukt for å generere det tilfeldige signalet opprinnelig (husk hva vi så av beat-frekvenser i høyre del av figur 3.)

Forslag til parametre:

Bruk et lavpassfilter som bare slipper gjennom frekvenskomponenter under 1000 Hz. Anvend dette filteret for å beregne signalet som kommer ut av hhv detektor A og B før krysskorrelasjonen bestemmes.

DEL 3: HBT I STØRRE DETALJ

Med alle detaljer vi nå har gått gjennom, ligger det til rette for å modellere HBT-effekten fullt ut. For at ikke oppgaven skal bli for omfattende, vil vi begrense oss til relativt få kilder og få detaljer mhp deteksjon, men det ligger i kortene hvordan vi kan utvide detaljrikdommen dersom vi skulle ønske det.



Figur 8: Geometrien vi skal bruke for å modellere HBT-effekten ved bruk av 11 støykilder (lyd) og to detektorer (mikrofoner).

Utgangspunktet er at vi har en kilde (f.eks. stjernen Sirius) som sender ut mange tilfeldige signaler, og det er ikke noe direkte sammenheng mellom frekvens, amplituder, faser og tidsforløp for lys som kommer fra en del av stjernen og lys fra en annen del av stjernen. Vi skal imidlertid ikke modellere lys, men lyd, og lar kilden bestå av 11 kilder som ligger med samme avstand fra hverandre langs en rett linje over en total avstand s (se figur 8). Kildene gir fra seg tilfeldig støy, og vi velger å la denne ha en gaussisk frekvensfordeling med senterfrekvens 7 kHz og helverdibredde på 6 kHz. Det er forskjellig signal fra hver av de 11 kildene.

Det er satt opp to detektorer (mikrofoner) A og B langs en linje parallell med linjen gjennom lydkildene. Avstanden mellom detektorene er a . Avstanden mellom de to linjene (vinkelrett på de to linjene) er h .

La normalen fra detektor A inn på linjen gjennom lydkildene treffe midten på kildearrayen (som vist på figur 8). Programmér en funksjon som beregner avstanden mellom A og hver av de 11 lydkildene og avstanden mellom B og hver av de 11 lydkildene. Velg noen konkrete verdier for s , h og a og sjekk ved egnede stikkprøver at programmet ditt gir samme resultat

som når du gjør en direkte beregning med kalkulator. Angi kort i rapporten din hvordan du gjennomførte testen.

Poenget er nå at lyd fra de 11 kildene når både detektor A og B, men gangveien fra hver enkelt kilde til hver av detektorene er litt forskjellig. Det betyr at vi får en summasjon av lydbølger både ved detektor A og B, men at summasjonen i prinsippet kan bli forskjellig. Det betyr at vi både for detektor A og B må legge sammen signalene fra de 11 kildene med i prinsippet ulik tidsforsinkelse for hvert bidrag - på en lignende måte som i forrige deloppgave.

Hvor stor blir forsinkelsen? Den bestemmes av forskjellen i veilengde δr for lyden sammen med lyd hastigheten i luft v . Vi velger referanseavstand lik avstanden mellom linjene hvor kildene er og linjen der detektorene er. Vis at for våre beregninger er tidsforsinkelsen gitt som

$$S_k(n) \rightarrow S_k(n + \delta n)$$

hvor S_k er arrayen som beskriver bølgen fra kilde k , og $\delta n = \text{round}(F_s \cdot \delta r / v)$ hvor F_s er samplingsfrekvensen.

Spesifikasjon av hvordan programmet skal fungere:

- Lag et dataprogram som generer 11 forskjellige tilfeldige lyd signaler, alle med en gaussisk frekvensfordeling med senterfrekvens 7 kHz og helverdibredde på 6 kHz. Summer lyd signalene (med riktig tidsforskyvning) fra disse lydkildene på detektor A's sted. Gjør det samme for detektor B for ett valg av a (detektoravstand).
- Beregn intensiteten for innsignalene til de to detektorene (kvadratet av utslaget), og anvend et lavpassfilter med grensefrekvens 1000 Hz på hvert av de to signalene for å simulere signalet som kommer ut av detektorene. Vi kaller disse signalene for hhv A2LP og B2LP.
- Beregn kryss-korrelasjonen mellom signalene i A2LP og B2LP.
- Velg etter tur en rekke detektoravstander a og gjennomfør alle beregningene på ny. Lag et (grovt) plot som viser hvordan krysskorrelasjonen avtar mot et plåtå (ofte med betydelig støy) når avstanden a blir stor.

Spesifikasjon av parametre for signaler som skal studeres:

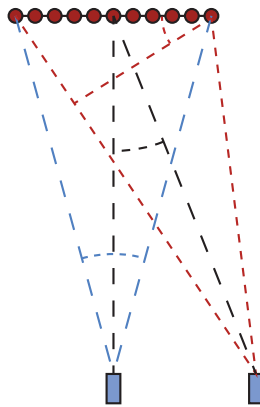
Følgende parametre anbefales brukt:

$s = 10$ m, $h = 200$ m, a skal varieres fra 0 til f.eks. 5 m, men tettest med punkter for små a (step på 5 cm i starten?).

Gjenta beregningene også for $h = 100$ m og 400 m.

Resultater vi er mest interessert i:

1. Beregn størrelsen $\eta = \theta a/\lambda$ for den verdien av a hvor krysskorrelasjonen har sunket halvveis ned mot platåverdien for alle de tre tilfellene av h . Finner du en viss lovmessighet?
2. Vis at størrelsen η har et lignende opphav som ligningen $d \sin \theta = \lambda$ i teorien for interferens fra en dobbeltspalt (eller optisk gitter eller diffraksjon fra en enkeltspalt). Figur 9 viser en skisse som muligens kan være til hjelp ved utledningen av relasjonen $\eta = \theta a/\lambda$.
3. For små a vil du forhåpentligvis finne at krysskorrelasjonen ikke endrer seg overhodet (eller i bare ett hakk). Det betyr at det er svært liten forskjell mellom signalet på detektor A og B (iblant ingen forskjell pga endelig tidsopløsning i vår modellering). Hva vil dette egentlig si mhp bølgefront og koherens på tvers av bølgens utbredelsesretning (romlig koherens)?



Figur 9: *Tips til et utgangspunkt som kan brukes for å vise relasjonen $\eta = \theta a/\lambda$.*

DEL 4: ANVENDER LOVMESSIGHETER FOR AN- DRE FYSISKE SITUASJONER

Nedenfor er det gitt fem ulike arbeidsoppgaver hvor man tar utgangspunkt i det vi har gjennomgått i denne prosjektoppgaven. Vi ønsker at du besvarer to av disse fem deloppgavene. Velg selv hvilke.

- I HBT-eksperimentet betraktet de intensitetskorrelasjoner mellom to detektorer for lys som var plassert flere meter fra hverandre. De kunne anslå at vinkelen tvers over stjernen Sirius, sett fra Jorda, måtte være om lag $0.0068''$ (0.0068 buesekunder!). Observasjonene (korrelasjonskoeffisienten vs avstand mellom A og B) er vist i figur 10. Er deres resultat omtrent som forventet i forhold til hva du har funnet for den forenklete modelleringen vi har brukt her (for lyd)? Huske at forskjellen i frekvens mellom HBT og vår modellering med lyd er om lag en faktor 10^{13} !

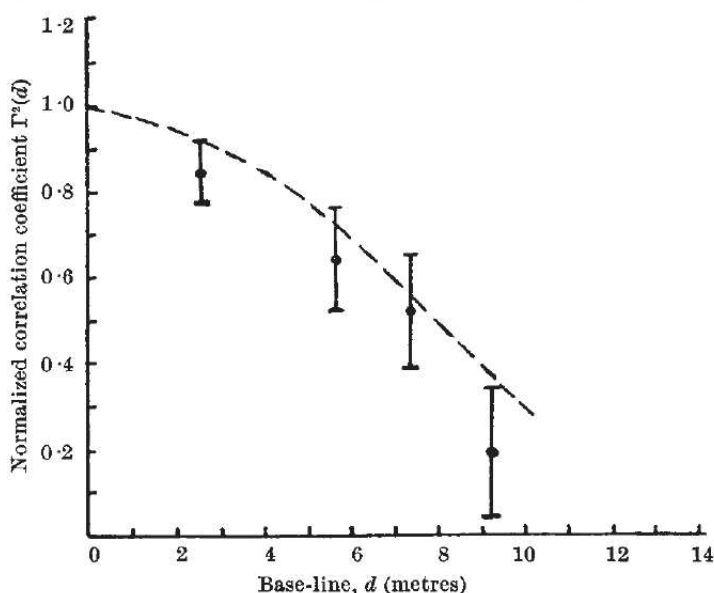


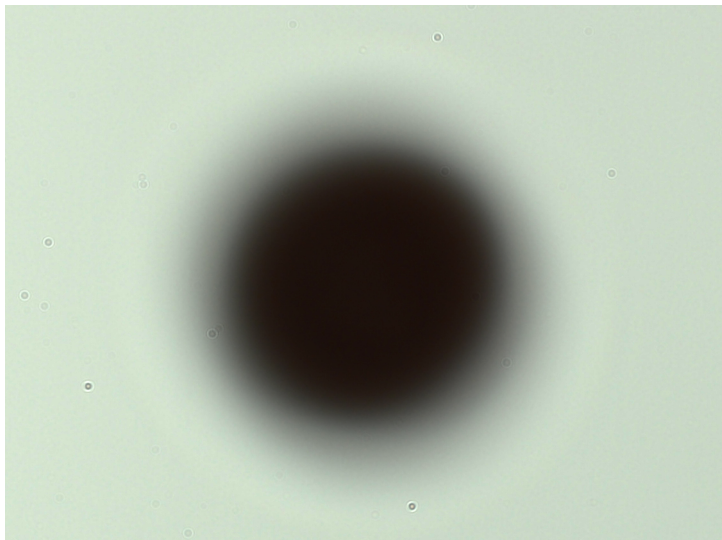
Fig. 2. Comparison between the values of the normalized correlation coefficient $\Gamma^2(d)$ observed from Sirius and the theoretical values for a star of angular diameter $0.0063''$. The errors shown are the probable errors of the observations

Figur 10: Originalfiguren hentet fra Hanbury Brown og Twiss' klassiske artikkel fra 1956 (se referanselisten).

- I vår moderne tid er det enkelt å vise interferens når vi sender lys gjennom en dobbeltspalt, fordi vi har lasere tilgjengelig som har tilnærmet perfekte plane bølgefronter på tvers av laserstrålen (under visse betingelser). Det er atskillig vanskeligere å vise interferens ved å bruke f.eks. sollys eller lampelys, som var det Thomas Young hadde tilgjengelig for sitt dobbeltspalteksperiment i 1801, og som Arago hadde tilgjengelig da han oppdaget "Aragos flekk" i skyggebildet av en kule tidlig på 1820-tallet.

Beregn bredden på de vel definerte bølgefrontene på lys direkte fra Sola (ved Jordens overflate). Dette svarer til den såkalte romlige koherensen til bølgene.]

- Figur 11 viser et fotografi tatt av skyggebildet til en metallkule med diameter 3.0 mm i direkte sollys. Kula var festet til en glassplate. Bildet ble samlet direkte opp på bildebrikken i et fotoapparat, og har ikke gått gjennom noe objektiv. Vi ser ikke noe til Aragos flekk i fotografiet, og den lyse flekken kom ikke fram for noen av de testede avstandene mellom kamera og kule. Derimot ser vi tydelige sirkler som har oppstått rundt skyggebildet til støvpartikler som fantes på glassplaten hvor kula var festet. Kan du forklare det vi observerer?



Figur 11: *Fotografi av skyggebildet av en 3.0 mm Ø metallkule i direkte sollys. Merk de bitte små sirklene rundt støvkornene på den tynne glassplaten hvor kula var festet.*

Kan du foreslå en metode for å kunne vise både interferens fra en dobbeltspalt og Aragos flekk ved hjelp av Sollys?

[Dersom du er interessert, kan vi dersom tid og vær og veilederpågang tillater det forsøke å lage et oppsett for å teste om det du kommer fram til faktisk fungerer i praksis.]

- Sammenligning med laser speckles
Laser speckles er et fenomen som har en viss analogi til HBT-fenomenet, men som også er ganske forskjellig. Laser speckles er jo et fenomen som først ble kjent etter at laseren ble oppfunnet i 1960. Finn fram til egnet beskrivelse av fenomenet og skriv din egen forklaring og få fram likheter og ulikheter med HBT-fenomenet. Beskrivelsen bør være på mellom en og to A4-sider. (Husk at du bare kan bruke andres tekst/formuleringer direkte dersom du markere dem som sitater og angir kilde.)

- Forklaringen på HBT-effekten har vært kontroversiell i alle år, noe som kommer delvis fram av en PowerPoint presentasjon til Indianara Lima Silva i 2010 (se referanselisten). Noen mener at “den korrekte forklaringen” er Glaubers innviklede forklaring basert på kvantestatistikk og kvantekorrelasjoner blant fotoner. Les Silva’s presentasjon og forsøk å gi et kort resyme av hva kontroversen gjelder.

DEL 5: OPPSUMMERING

Helt til slutt i prosjektoppgaven din ber vi deg gi en oppsummering av fysikken som ligger bak HBT-effekten der vi håper du klarer å bruke resultater fra store deler av prosjektoppgaven. Klarhet og faglig god fremstilling bør etterstrebnes. Oppsummeringen bør være mellom en halv og en hel A4-side. Denne delen av rapporten tillegges stor vekt ved bedømmingen av besvarelsen.

Det anbefales at du diskuterer med medstudenter hvordan stoffet kan presenteres. Du kan også gjerne bryne deg mot kursets lærere, men bare fram til og med fredagen. Husk at det er du selv som må ta initiativet!

REFERANSER

R. Hanbury Brown, R.Q. Twiss: A test of a new type of stellar interferometer on Sirius. Nature 178 (1956) 1046 - 1048.

[Dersom du leser denne artikkelen og ønsker å hente opp artikkelen forfatterne henviser til når det gjelder matematisk analyse (artikkel i Phil. Mag. 45 (1954) 663 - 682), er ikke den artikkelen tilgjengelig elektronisk. Send Arnt Inge en mail dersom du er interessert i artikkelen.]

Wikipedia (eng), under oppslagsord: Hanbury Brown and Twiss effect.

Indianara Lima Silva: The controversy of photons and the Hanbury Brown and Twiss experiment. HQ-3 Conference on the history of quantum physics, Berlin, July 2010. Tilgjengelig 14. april 2014 på

http://quantum-history.mpiwg-berlin.mpg.de/news/workshops/hq3/hq3_talks/23_silva.pdf.

VEDLEGG: FORKLARING PÅ ENKELTE RELEVANTE ORD

Vi har sakset litt fra Wikipedia for å gi en forklaring på en del aktuelle ord og uttrykk i denne prosjektoppgaven. Vi beholder beskrivelsene på engelsk. Iblant har vi justert litt på setningene for at de skal kunne brukes i kortform. God fornøyelse.

STOKASTISK PROSESS

In probability theory, a stochastic process, or sometimes random process (widely used) is a collection of random variables; this is often used to represent the evolution of some random value, or system, over time. This is the probabilistic counterpart to a deterministic process (or deterministic system). Instead of describing a process which can only evolve in one way (as in the case, for example, of solutions of an ordinary differential equation), in a stochastic or random process there is some indeterminacy: even if the initial condition (or starting point) is known, there are several (often infinitely many) directions in which the process may evolve.

STASJONÆR

In mathematics, a stationary process (or strict(ly) stationary process or strong(ly) stationary process) is a stochastic process whose joint probability distribution does not change when shifted in time. Consequently, parameters such as the mean and variance, if they are present, also do not change over time and do not follow any trends.

ERGODISK

A stochastic process is said to be ergodic if its statistical properties (such as its mean and variance) can be deduced from a single, sufficiently long sample (realization) of the process.

AUTOKORRELASJON

Given a signal $f(t)$, the continuous autocorrelation $R_{ff}(\tau)$ is most often defined as the continuous cross-correlation integral of $f(t)$ with itself, at lag τ .

$$R_{ff}(\tau) = (f(t) \cdot \bar{f}(-t))(\tau) = \int_{-\infty}^{\infty} f(t + \tau) \bar{f}(t) dt = \int_{-\infty}^{\infty} f(t) \bar{f}(t - \tau) dt$$

where \bar{f} represents the complex conjugate and $*$ represents convolution. For a real function, $\bar{f} = f$.

The discrete autocorrelation R at lag j for a discrete signal x_n is

$$R_{xx}(j) = \sum_n x_n \bar{x}_{n-j}.$$

...

For processes that are also ergodic, the expectation can be replaced by the limit of a time average. The autocorrelation of an ergodic process is sometimes defined as or equated to [1]

$$R_{ff}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(t + \tau) \bar{f}(t) dt$$

$$R_{xx}(j) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} x_n \bar{x}_{n-j}$$

. These definitions have the advantage that they give sensible well-defined single-parameter results for periodic functions, even when those functions are not the output of stationary ergodic processes.

Alternatively, signals that last forever can be treated by a short-time autocorrelation function analysis, using finite time integrals. (See short-time Fourier transform for a related process.)

CORRELATION FUNCTION C

A correlation function is a statistical correlation between random variables at two different points in space or time, usually as a function of the spatial or temporal distance between the points. If one considers the correlation function between random variables representing the same quantity measured at two different points then this is often referred to as an autocorrelation function being made up of autocorrelations. Correlation functions of different random variables are sometimes called cross correlation functions to emphasise that different variables are being considered and because they are made up of cross correlations.

Correlation functions are a useful indicator of dependencies as a function of distance in time or space, and they can be used to assess the distance required between sample points for the values to be effectively uncorrelated. In addition, they can form the basis of rules for interpolating values at points for which there are observations.

CROSS-CORRELATION FUNCTION

In signal processing, cross-correlation is a measure of similarity of two waveforms as a function of a time-lag applied to one of them. This is also known as a sliding dot product or sliding inner-product. It is commonly used for searching a long signal for a shorter, known feature. It has applications in pattern recognition, single particle analysis, electron tomographic averaging, cryptanalysis, and neurophysiology.

For continuous functions f and g , the cross-correlation is defined as:

$$(f \star g)(\tau) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f^*(t) g(t + \tau) dt$$

where f^* denotes the complex conjugate of f and t is the time lag.

Similarly, for discrete functions, the cross-correlation is defined as:

$$(f \star g)[n] \stackrel{\text{def}}{=} \sum_{m=-\infty}^{\infty} f^*[m] g[m + n].$$

Visual comparison of convolution, cross-correlation and autocorrelation. The cross-correlation is similar in nature to the convolution of two functions.

In an autocorrelation, which is the cross-correlation of a signal with itself, there will always be a peak at a lag of zero unless the signal is a trivial zero signal.

GAUSSISK FORDELING

In probability theory, the normal (or Gaussian) distribution is a very commonly occurring continuous probability distribution - a function that tells the probability that an observation in some context will fall between any two real numbers.

A normal distribution is

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

The parameter μ in this definition is the mean or expectation of the distribution (and also its median and mode). The parameter σ is its standard deviation; its variance is therefore 2σ . A random variable with a Gaussian distribution is said to be normally distributed and is called a normal deviate.

Programsnutter (Matlab)

Først gir vi en liten snutt som kanskje kan være til hjelp når du skal generere et stokastisk (random) signal med en gaussisk amplitudfordeling i frekvensspekteret og vilkårlige faser. Det er bare lagt opp til generering av nedre halvdel av frekvensspekteret. Man må generere øvre halvdel av frekvensspekteret ved å basere seg på kunnskap om folding gitt i kapittel 4 i læreboka.

Dernest er det gitt et eksempel på organisering av et hovedprogram. Vi kaller de ulike funksjonene etter som vi trenger dem. Merk at vi forventer at det gis korrekte verdier og tekst langs aksene på de plottene som brukes i rapporten du lager.

```

y = zeros(N,1);
T = N/Fs;
t = linspace(0,T*(N-1)/N,N);
f = linspace(0,Fs*(N-1)/N, N);
nsenter = floor(N*fsenter/(Fs*(N-1)/N));
nsigma = floor(N*fsigma/(Fs*(N-1)/N));

gauss = exp(-(f-fsenter).*(f-fsenter)/(fsigma*fsigma));
% figure;
% plot(f,gauss,'-k'); % For sjekking!
ampl = rand(N,1);
ampl = ampl.*transpose(gauss);
% figure;
% plot(f,ampl,'-g'); % For sjekk
faser = rand(N,1);
faser = faser*2*pi;
y = ampl.*(cos(faser) + i* sin(faser));

```

```

function TestDell

\% Lager et tilfeldig signal som ønsket
Fs = 44100;
N = 1024*64;
T = N/Fs;
t = linspace(0,T*(N-1)/N, N);
f = linspace(0,Fs*(N-1)/N,N);
fsenter = 10000;
fsigma = 5000/2;
xx = HvitStoyGauss(Fs,N,fsenter,fsigma);

\% Velger å plotte tids- og frekvensbildet her
figure;
plot(t,xx,'-g');
xlabel('Tid (s)');
ylabel('Signal (vilkaarlig enhet)');

figure;
plot(f,abs(fft(xx)),'-r');
xlabel('Frekvens (Hz)');
ylabel('Frekvenskomponent (rel. verdier)');

\% Wavelet analyse
fmin = 7000.0; \% Minimum frekvens i waveletanalysen (i Hz)
fmax = 13000.0; \% Maximum frekvens
K = 48; \% Morlet-wavelet-bredde (kan være 6 - 400+)
WL1(xx,N,fmin,fmax,K,Fs); \% Kaller på wavelettransformasjon

\% Autokorrelasjons-analyse
AutoCorr(xx,N,Fs);

```

Liste over funksjoner/programmer det legges opp til at du bør lage:

1. *HvitStoyGauss* : Funksjon for å generere et stokastisk signal med gitt gaussisk frekvensfordeling.
2. *WaveletTransform* : Funksjon som beregner den wavelet-transformerte av et signal og plotter resultatet.
3. *AutoCorr* : Funksjon som beregner og plotter autokorrelasjonsfunksjonen til (halvparten av) et signal. Det kan iblant være vel så lurt å plote absoluttverdien til autokorrelasjonsfunksjonen.
4. *CrossCorr* : Funksjon som beregner krysskorrelasjonen for to signaler.
5. *LowPass* : Funksjon som filtrerer bort alle frekvenskomponenter med frekvens over en gitt grense.
6. *Avstander* : Funksjon som regner ut alle kombinasjoner av avstander mellom kilder og detektorer i del 2b.
7. *Del1Prog* : “Hovedprogram” hvor du får gjort det meste i del 1 av prosjektoppgaven. Kaller på noen av funksjonene ovenfor.
8. *Del2aProg* : Tilsvarende “hovedprogram” for del 2a av prosjektoppgaven.
9. *Del2bProg* : Tilsvarende for del 2b.

Du står selvfølgelig fritt i å legge opp programmeringen som du vil, - men programmene må testes og fungere som de skal, og koden bør være rimelig lett å lese og forstå. Programmene legges som vedlegg til rapporten din.