



15. Koherens, dipolstråling og laser

Dette kapitlet tar opp følgende temaer: Koherens - kvalitativt og matematisk, (auto)korrelasjonsfunksjoner, syntese og analyse av vilkårlige signaler, anskueliggjøring av koherens, måling av koherenslengder for lys, stråling fra en elektrisk ladning, dipolstråling, lasere. Lesetips finnes på websiden med hjelpefiler til boka.

15.1 Koherens, en kvalitativt tilnærming

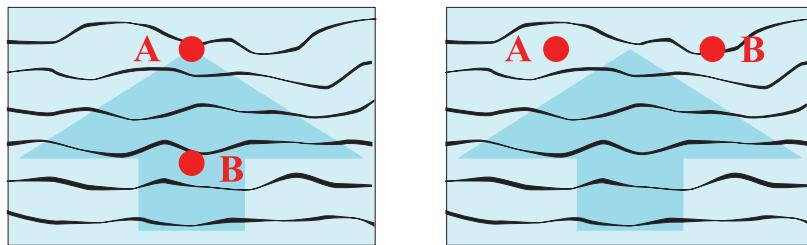
“Koherens” er et ord, uttrykk eller begrep som få nordmenn har et forhold til. Ordet er antakelig brukt mer i det engelske språk, og i en dictionary kan vi finne forklaringer på “coherence” så som: “The quality of being logically integrated, consistent, and intelligible”, eller: “The mutual relationship between sets of electromagnetic or sound waves in which their amplitudes are exactly equivalent and rise and fall together.”

Koherens er et begrep knyttet til bølger eller tidsvariable signaler. Når bølger på to punkter A og B i rommet er koherente, mener vi at det er en bestemt relasjon mellom bølgene som til enhver tid passerer punktene A og B. Et eksempel på en slik relasjon er at bølgene alltid har sine maksimumsutslag samtidig i A og B (eller at det f.eks. er κT tidsforskjell mellom tidene der bølgen har sitt maksimalutslag i A og B, der T er en midlere periodetid og κ er en reell konstant).

Vi skiller mellom tidmessig og romlig koherens (engelsk: temporal & spatial coherence). For **tidmessig koherens** vurderer vi bølgen i to punkter A og B som ligger etter hverandre i forhold til bølgens bevegelsesretning (se figur 15.1). Vi sjekker da om det er en vedvarende bestemt relasjon mellom bølgen i punktet A med bølgen i et annet punkt B som samme del av bølgen passerte for litt siden.

For **romlig koherens** sammenligner vi bølgen på et sted A med bølgen på et sted B som ligger ved siden av hverandre i forhold til bølgens bevegelsesretning.

Koherens er en viktig betingelse for å få interferens. Ved utledning av intensitetsfordelingen etter en dobbelspalt antok vi at bølgen hadde samme utslag til enhver tid i de to spalteåpningsene. Det er ekvivalent med å si at den romlige koherenslengden må være minst så lang at den dekker begge spaltene samtidig.



Figur 15.1: *Tidmessig og romlig koherens forteller oss noe om regelmessigheten i bølger hhv i den retningen bølgen beveger seg og i en retning vinkelrett på denne. Bølgen antas å være todimensjonal i dette tilfellet (f.eks. overflatebølger på vann). De sorte stripene angir bølgetopper, og tykkelsen angir amplituden bølgen har på det aktuelle stedet. Bølgen i denne figuren er temmelig uregelmessig.*

Vi har hittil beskrevet bølger på en matematisk, idealisert måte, nemlig som:

$$f(x, t) = A \cos(kx - \omega t + \phi) \quad (15.1)$$

Ingen virkelige bølger følger et slikt matematisk perfekt forløp. I figur 15.1 er bølgen nokså uregelmessig. I venstre del av figuren, ved den tiden figuren angir, har tilfeldigvis en bølgetopp nettopp passert både punkt A og B. Litt senere vil det være en bølgedal samtidig i punkt A og B. Dersom bølgen har den egenskapen at hver gang vi har en bølgetopp i A, så vet vi hvilken fase bølgen har i punkt B, er det tidmessig koherens i punkt A og B.

Dersom det derimot f.eks. er slik at når vi vet at vi har en bølgetopp i A, vet vi likevel ikke hvilken fase det er i bølgen i punkt B, er det dårlig korrelasjonen mellom bølgen i punkt A og B. Det er i så fall ingen tidmessig koherens mellom bølgen i punktene A og B.

På samme måte kan vi vurdere bølger i to punkter som ligger *ved siden av hverandre* for å bestemme grad av romlig koherens.

For virkelige bølger er det alltid høy grad av korrelasjon mellom bølgene i A og B dersom de to punktene er svært nær hverandre (mindre enn den minste bølgelengden som finnes i bølgen). Derimot er det alltid slik for reelle bølger at korrelasjonen mellom bølgene i A og B blir elendig dersom avstanden mellom A og B gjøres stor nok.

Dersom avstanden mellom A og B må være mindre enn noen få bølgelengder for å få høy grad av korrelasjon, sier vi at bølgen er *ikke koherent*. Dersom vi derimot kan finne høy grad av korrelasjon også når avstanden mellom A og B er svært mange bølgelengder, sier vi at bølgen *koherent*. Det er likevel en kontinuerlig overgang mellom ikke-koherens og koherens.

Når sammenligningen skjer i punkter *langs* retningen bølgen brer seg eller *vinkelrett på*, er det hhv snakk om *tidmessig* eller *romlig* koherens.

Den største avstanden vi kan ha mellom A og B som fortsatt gir betydelig korrelasjon, kaller vi *koherenslengden*. Vi kan angi både en tidmessig og en romlig koherenslengde. For tidmessig koherens kan vi også operere med *koherenstid* som er den tiden bølgen bruker på å vandre den tidmessige koherenslengden.

Siden det alltid er en viss grad av variabilitet i virkelige bølger, og variabilitet kan beskrives ved statistikk, er koherens en statistisk størrelse.

15.1.1 Matematisk håndtering av koherens

Skal vi karakterisere en bølge, kan vi enten registrere utslag på ett sted i rommet som funksjon av tid, eller vi kan registrere utslag ved ett tidspunkt som funksjon av posisjon i rommet. I begge tilfeller ender vi i virkelige målinger opp med en tallrekke, en array. Gjør vi målinger i to punkter, slik det legges opp til i figur 15.1, ender vi opp med to tallrekker, to signaler.

Det er mange måter man kan sammenligne to slike tallrekker på, men ved analyse av koherens har fysikere for mange år siden valgt å bruke en liknende angrepsmåte som fouriertransformasjon. Ved fouriertransformasjon foretar vi egentlig en korrelasjonsanalyse mellom det signalet vi studerer og en perfekt matematisk harmonisk funksjon med en gitt frekvens (og vi endrer på frekvensen for å få hele frekvensspekteret). Fouriertransformasjonen innebærer altså beregning av indreproduktet mellom signalet vi analyserer og en perfekt harmonisk funksjon.

Ved analyse av korrelasjon mellom signaler (bølger), regner vi også ut indreproduktet, men nå mellom de to signalene vi sammenligner! På en måte bruker vi det ene signalet som referanse for å sjekke hvor mye det ligner på det andre signalet.

Dersom signalet vi registerer i punkt A kalles $f(t)$ og signalet i B kalles $g(t)$, har folk før oss valgt å beregne korrelasjon mellom f og g på følgende måte:

$$C = \frac{\int f(t)g(t)dt}{\sqrt{\int f^2(t)dt \int g^2(t)dt}} \quad (15.2)$$

Integrasjonene går over en litt vilkårlig tidslengde. Er signalene stasjonære (endrer ikke statistisk karakter over tid) og atpå til ergodiske (at all statistikk vi kan få ved å gjøre analyser på mange uavhengige signaler er lik statistikken vi oppnår dersom vi nøyer oss med å følge bare ett signal lenge nok), vil C gå mot en veldefinert verdi når integrasjonstiden økes.

Korrelasjon definert på denne måten er rett og slett ett tall. Det blir ikke noe funksjon før vi på et eller annet systematisk vis endrer på hvordan f og g måles eller genereres.

Autokorrelasjon

I venstre del av figur 15.1 valgte vi å sammenligne bølger ved to punkter A og B som lå etter hverandre i den retningen som bølgen beveger seg. Praktiske målinger kan gjennomføres temmelig uproblematisk dersom vi skal analysere lyd, for vi kan lage meget små mikrofoner som ikke forstyrrer bølgen så mye at signalet i A blir endret ved nærværet av mikrofonen i punkt B.

I andre sammenhenger er det mer problematisk å legge inn en sensor i B uten at bølgen som når punkt A blir forstyrret.

Det er en av grunnene til at når vi analyserer tidmessig kohärens for en bølge, velger vi ofte å bare bruke én detektor, f.eks. i punkt B, og dropper helt den i A. Antar vi at formen på det overordnede bølgemönsteret (som gitt ved de sorte strekene i figur 15.1) ikke endrer seg mye i løpet av tiden det tar for bølgen å bevege seg en avstand fra B til A, vil signalet i punkt A være tilnærmet det samme som i punkt B, bare tidsforskjøvet (på grunn av at bølgemönsteret beveger seg med en gitt hastighet).

I så fall vil

$$f(t) \approx g(t + \tau)$$

hvor τ er tiden bølgen (bølgemönsteret) bruker på å bevege seg avstanden fra B til A.

Den tidmessige korrelasjonen mellom bølgen i A og B er da gitt ved:

$$C = \frac{\int g(t)g(t + \tau)dt}{\sqrt{\int g^2(t)dt \int g^2(t + \tau)dt}} \quad (15.3)$$

Dersom tidsvariasjonen ikke endrer karakter over tid sier vi at bølgen er *stasjonær*. For stasjonære bølger vil

$$\int g^2(t)dt \approx \int g^2(t + \tau)dt$$

Med denne betraktingen ligger det til rette for at vi selv med bare én måleserie kan beregne korrelasjoner for mange ulike avstander mellom punktene A og B. I praksis gjør vi dette ved å endre tidsforskyvningen τ i ligning (15.3) overfor. Lar vi τ i ligning (15.3) variere kontinuerlig fra null og et stykke framover, vil korrelasjonene $C(\tau)$ bli til det vi kaller *autokorrelasjonsfunksjonen* for signalet.

Uttrykket for autokorrelasjonsfunksjonen for et signal g blir da:

$$C(\tau) = \frac{\int g(t)g(t + \tau)dt}{\int g^2(t)dt} \quad (15.4)$$

For et digitalisert signal $g(i)$ for $i = 1, \dots, N$ (diskrete tidspunkt), vil uttrykket se slik ut:

$$C(j+1) = \frac{\sum_{i=1}^M g(i)g(i+j)}{\sum_{i=1}^M g(i)g(i)} \quad (15.5)$$

for $j = 0, \dots, N - M$.

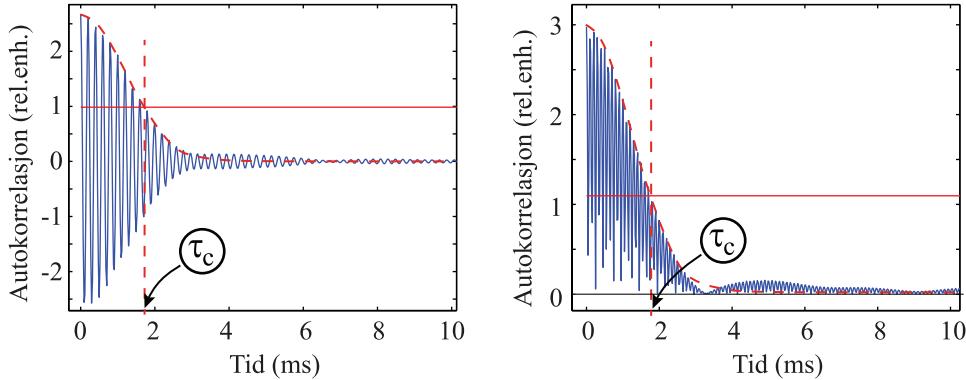
Merk at siden vi forskyver utvalget av punkter som brukes for å beskrive signalene i punkt A og B ut fra samme tidsstrek av signalet, må $M < N$. Den største j -verdien vi kan beregne for et signal som er beskrevet i N punkter, er $j = N/2$.

Venstre del av figur 15.2 viser et eksempel på en autokorrelasjonsfunksjon for en bølge. Langs x-aksen har vi tidsforskjellen som $f(t) = g(t + \tau)$ er forskjøvet i forhold til $g(t)$. Dersom

samplingsfrekvensen er F_s vil sammenhengen mellom indeksen j og tidsforskyvingen τ være:

$$j = \text{round}(F_s\tau)$$

hvor "round" står for at man finner nærmeste heltallsverdi for uttrykket i parentesene.



Figur 15.2: Et eksempel på hvordan autokorrelasjonsfunksjonen til en bølge kan se ut. Til venstre er den vanlige autokorrelasjonsfunksjonen plottet, mens i høyre del er absoluttverdien plottet. Tidsforskyvningen τ_c som fører til at korrelasjonen har sunket til $1/e$ av maks verdi, kaller vi bølgens koherenstid.

Vi ser at autokorrelasjonen varierer fra +1 til nær -1, og oscillerer opp og ned samtidig som amplituden i variasjonen avtar mot null. Korrelasjon lik 1 (for normert formel slik vi har brukt her), svarer til at f og g er identiske. Det er de alltid i første punktet (ingen forskyving). Korrelasjon -1 ville bety at $f = -g$. For en bølge ville det bare si at når f har en bølgetopp, vil g ha en bølgebunn og omvendt. I så fall er det fortsatt en perfekt korrelasjon, bare at den er nettopp fortegnsskiftet.

Når avstanden mellom punkt A og B i figur 15.1 øker, blir det i figur 15.2 en gradvis overgang fra høy korrelasjon (øvre tenkte omhyllingskurve har en verdi nær 1) til en mindre og mindre korrelasjon (omhyllingskurven nær null). Etter ca 1.8 ms har tallverdien sunket til $1/e$ av max. Vi sier da at den tidmessige koherenstiden τ_c for dette signalet er 1.8 ms. Den tilsvarende tidmessige koherenslengden er $\tau_c v$ hvor v er bølgens fasehastighet.

I høyre del av figur 15.2 har vi plottet tallverdien av autokorrelasjonsfunksjonen. Vi har jo sett at korrelasjon +1 og -1 begge svarer til en perfekt korrelasjon, bare med et fortegnsskifte i bølgeutslaget i det siste tilfellet. Det er derfor best å trekke en tenkt omhyllingskurve gjennom toppunktene i absoluttverdien til autokorrelasjonsfunksjonen når vi skal bestemme hvilken forskyvning som svarer til at korrelasjonen har sunket fra 1 til $1/e$.

[♣ ⇒ Det bør bemerknes at i praksis vil aldri autokorrelasjonen gå helt mot null når avstanden mellom A og B øker. Imidlertid, dersom gjennomsnittsverdien for g er null, og dersom strengen er lang nok i forhold til variasjoene i signalet, vil autokorrelasjonen for store τ komme så nær null at den resterende variasjonen ikke påvirker bestemmelsen av koherenstiden nevneverdig.]

Det viser seg at variasjonen i første del av autokorrelasjonsfunksjonen er temmelig stabil dersom vi gjør flere etterfølgende opptak av signalet $g(t)$, men oscillasjonene rundt null når vi har passert minst to ganger koherenstiden vil endre seg fra dataopptak til dataopptak.

Tar vi gjennomsnittet av mange tidsforløp og legger sammen autokorrelasjonene, vil den første delen av korrelasjonene adderes konstruktivt mens det vi gjerne kaller "støyen" omkring null etter hvert vil forsvinne. Autokorrelasjonsfunksjonen kan da ofte tilnærmet skrives:

$$C(\tau) = \cos(\bar{\omega}\tau)e^{-(\frac{\tau}{\tau_c})^2} \quad (15.6)$$

hvor $\bar{\omega}$ svarer til middelverdien av vinkelfrekvensene i det opprinnelige signalet. τ_c er korrelasjonstiden og svarer (med et visst forbehold) til det vi også kaller "koherenstiden" for vår bølge/swingning. Vi kan med noen forbehold si at koherenstiden defineres ved ligning (15.6). $\Leftarrow \spadesuit$

Dersom g var en perfekt sinus, ville en τ_j som svarer til en forskyvning på $2\pi n$ (n er et heltall) i sinussignalet, medføre at $g(i)g(i+j)$ i ligning (15.5) i praksis er en \sin^2 -funksjon. Autokorrelasjonsfunksjonen C vil da rett og slett bli middelverdien av \sin^2 dividert med samme verdi, hvilket er 1.

På lignende måte kan vi vise at når τ_j svarer til $(2n+1)\pi$ i det perfekte sinussignalet g , vil beregningen innebære middelverdien av $-\sin^2$, og svaret ville bli -1. For en forskyvning på en kvart bølgelengde i forhold til de tilfellene vi allerede har nevt, vil C bli middelverdien av et $\sin \cdot \cos$ -uttrykk, som er lik null.

Autokorrelasjonsfunksjonen for en perfekt sinus vil derfor bli en regelmessig funksjon som varierer fra +1 til -1. Autokorrelasjonsfunksjonen får rett og slett en cosinusform og vil aldri få en amplitud som avtar mot null slik som i figur 15.2. Koherenstiden ville da bli uendelig.

Ingen reelle bølger har uendelig koherenstid, men dersom tidsstrengen vi har tilgjengelig for analyse ikke er mer enn to ganger koherenstiden lang, klarer vi ikke å bestemme koherenstiden.

15.1.2 Reelle fysiske signaler

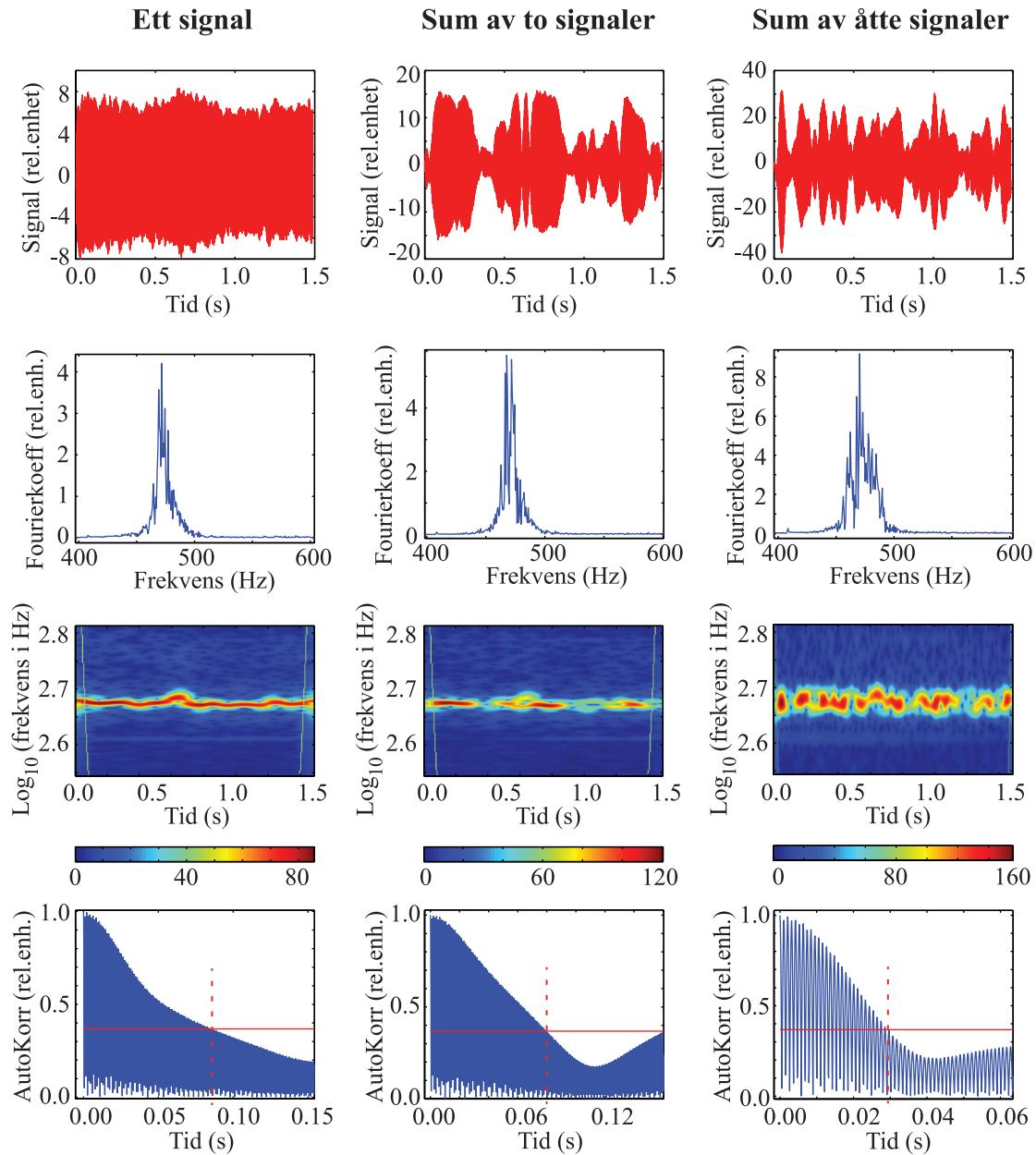
Vi skal nå ta for oss noen reelle fysiske signaler for å vise noen typiske trekk. Vi har tilgjengelig tre ulike analysemetoder for å studere tidssignalet, nemlig fouriertransformasjon, waveletanalyse og beregning av autokorrelasjonsfunksjonen.

Signalene er mikrofonsignaler etter sampling av lydbølger fra en person som synger "iiiii"-er hele tiden samlingen foregår. Det er foretatt åtte separate lydopptak av samme person. Fra disse opptakene har vi konstruert tre: Signal 1 er rett og slett ett av lydopptakene. Signal 2 er summen av to opptak, og signal 3 er summen av alle åtte opptakene. Signal 2 simulerer et dataopptak av to personer som synger 'iiiii"-er samtidig, og tilsvarende for signal 3. Signalene er vist i figur 15.3 sammen med analysene vi har foretatt.

Signal fra én kilde

Vi ser av venstre kollonne i figur 15.3 at amplituden i signalet (og dermed også lydintensiteten) holder seg temmelig lik i hele perioden. Frekvensspekteret har en bredde på om lag 15 Hz. I waveletdiagrammet kommer det fram hvordan forandringen i frekvens endret seg i løpet av dataopptaket. Nederst i figuren er den første delen av autokorrelasjonsfunksjonen gitt, og vi kan estimere koherenstiden for signalet til ca 0.18 s.

Lyd hastigheten i luft er om lag 340 m/s. Det betyr at vi kan forutsi korrelasjon i fasen i det aktuelle lydsignalet innenfor en strekning på ca $\Delta L = 340 \cdot 0.18$ meter, det vil si ca 60 m. Denne strekningen kaller vi *koherenslengden* for vår bølge (mer presist: *den tidmessige koherenslengden*, også kalt *den temporære koherenslengden*).



Figur 15.3: Tre eksempler på reelle samplede lydsignaler, med fra venstre til høyre én, to og åtte stemmer samtidig. Det er også vist frekvensanalyse (en liten annen harmonisk komponent er ikke tatt med), waveletanalyse og den første delen av autokorrelasjonsfunksjonen. For waveletanalysen ble $K=100$ brukt på de første to signalene, og $K=32$ på den siste.

Det synes ikke å være noe spesielt med dataopptaket for en kilde. Alt er som vi forventet, men vi innser at ustabiliteten i frekvens (tonehøyde) omtrent midt i dataopptaket nok påvirker den beregnede koherenstiden. Vi merker oss også at autokorrelasjonsfunksjonen ikke går pent og pyntelig ned mot null etter at vi har passert koherenstiden.

Signal fra flere kilder

Vi ser av den midterste kollonnen i figur 15.3 at amplituden i summen av to nokså like bølger varierer på en drastisk måte i løpet av dataopptaket, selv om amplituden i hvert av bidragene holdt seg temmelig stabilt. Grunnen til dette er at lydbølgene fra sanger 1 og sanger 2 i enkelte perioder adderes konstruktivt slik at amplituden blir omtrent dobbelt så stor som fra hver av dem. Samtidig vil lydbølgene i andre perioder adderes destruktivt, og amplituden på lydbølgene blir nær null.

Alle bølger synes å ha følgende egenskap:

Utslaget på et vilkårlig sted i bølgen kan ha opp til mange bidrag. Utslaget har imidlertid bare én verdi innenfor et tilstrekkelig lite volum (lengder i hver retning godt under minste bølgelengde i bølgene som bidrar). Eksempelvis vil lydtrykket bare ha én verdi i små volum når en lydbølge passerer, høyden på vannoverflaten har bare en verdi på ethvert sted, og det elektriske feltet i summen av alle elektromagnetiske bølger har bare en verdi i hvert bitte lite volum selv om mange elektromagnetiske bølger bidrar.

Vi kan bare gjenkjenne ulike bidrag til bølgene og hvor de kom fra ved å betrakte mønsteret i summen av alle bølgene og hvordan dette mønsteret utvikler seg i tid. Bidrag fra ringer i vann etter at vi har kastet en stein uti, kan bare gjenkjennes ved at vi får øye på ringene i området rundt det lille volumet vi betrakter.

Det er *sumbølgen* som utvikler seg i tid i henhold til bølgeligningen, ikke hvert enkelt bidrag. Når bølgeligningen er lineær kan vi likevel beskrive utviklingen av en bølge *matematisk* som en sum av flere bidrag som hver for seg utvikler seg i henhold til bølgeligningen. At vi kan bruke matematikk på den måten, bør imidlertid betraktes mer som et gunstig unntak heller enn hovedregelen (for når vi har med ikke-lineære prosesser å gjøre, gjelder det ikke).

Når vi adderer flere uavhengige bølger, blir det bestandig noen perioder med konstruktiv og noen perioder med destruktiv interferens. Varigheten av periodene med konstruktiv interferens avhenger svært mye av frekvensvariasjonen i de signalene som adderes. Vi vil vise flere eksempler på dette litt senere i kapitlet. Effekten kommer ofte vakkert fram i kontinuerlig waveletanalyse med Morlet wavelets.

Figur 15.3 viser foruten den karakteristiske fluktasjonen i amplitude på grunn av summasjonen, at frekvensspekteret får større bredde etter som flere signaler bidrar. Hvert signal har sin senterfrekvens og variasjon, og summen av signaler får derfor en større bredde enn hvert enkelt bidrag. I vårt tilfelle er det likevel ikke større forskjell enn 1 Hz i senterfrekvens for de åtte bidragene (testet separat, ikke vist).

En økning i bredden på frekvensfordelingen slår ut også i koherenstiden. Ved flere bidrag får vi en kortere koherenstid enn i enkeltbidragene. Det er imidlertid umulig å trekke slutninger om sammenhenger mellom f.eks. bredde på frekvensspekteret og koherenstid ut fra disse eksemplfilene alene. Statistikken er for dårlig til det. Vi kommer tilbake til problemstillingen om litt.

Waveletdiagrammene i kollonne 2 og 3 i figur 15.3 viser stort sett de samme karakteristiske trekkene som amplitudevariasjonen i tidsbildet. Frekvensen er imidlertid såpass godt definert at vi iblant kan se små endringer i den dominerende frekvensen i sumsignalet etter som tiden går.

Amplituden på sumsignalet

Det kan bemerkes at gjennomsnittsamplituden for summen av åtte lignende signaler er om lag 20 i en målestokk der gjennomsnittsamplituden til ett av signalene er om lag 7. Det kan lett vises at $20 \approx 7 \cdot \sqrt{8}$. Ved addisjon av uavhengige bølger med omtrent samme amplitude, frekvens og grad av variasjon med tiden, er ikke summen av amplituder proporsjonal med antall signaler som bidrar.

For harmoniske signaler med nær matematisk perfekt samme frekvens, amplitude og fase, er det derimot slik at summen får en amplitud som er proporsjonal med antall signaler vi adderer. Det er det som gjør at intensiteten i en laserstråle kan bli betydelig selv om effekten bare er noen få watt.

La oss gå tilbake til sangerne våre som synger “iiiii”-er. Dersom sangerne hadde sunget eksakt likt og hele tiden i fase, ville åtte sangere gitt en intensitet lik 64 ganger intensiteten til hver enkelt (intensitet er proporsjonal med amplitude kvadrert). Dette er ganske forskjellig fra 8 ganger intensiteten vi i praksis fant med våre reelle signaler. Er dette noe vi kan utnytte?

Imidlertid, “*There Ain’t No Such Thing As A Free Lunch*”. Vi får ingenting gratis. Det kan synes som om at åtte hypotetiske sangere som synger koherent lyd i fase ville gi *åtte ganger større* lydintensitet enn åtte sangere som ikke synger koherent. Hvordan kan da kravet om energibevaring tilfredsstilles?

Energiregnskapet går bra dersom vi tar i betraktning også *romlige* forhold. Åtte hypotetiske sangere som synger koherent vil ikke gi åtte ganger høyere lydenergien overalt i rommet, bare på de stedene der signalene fra alle åtte er i fase med hverandre. Andre steder i rommet, der signalene er i motfase, vil lydenergien kunne droppe til bortimot null. Slik vil det ikke være med de reelle sangerne. Det finnes ingen steder i rommet hvor det er permanent konstruktiv eller destruktiv interferens for disse sangerne. Vi vil høre lyden fra de reelle sangerne overalt.

Integreres lydenergien overalt i rommet, vil den bli omtrent den samme uansett om sangerne synger koherent eller ikke-koherent.

Disse romlige betraktningene har paralleller med intensitetsfordelingen for stripene fra flere spalter. Når antall spalter øker, blir stripene smalere og smalere, og stripene blir mer intense selv om totalt lysstrøm ut fra spaltene er uforandret.

15.2 Syntese og analyse av vilkårlige signaler

I eksemplet ovenfor brukte vi reelle lydopptak fra en sanger som sang “iiiiiiii”-er mens tonehøyde og amplitude ble forsøkt holdt konstant. Vi skal nå analysere signaler der vi selv kan bestemme variasjonen i bølgene på en annerledes måte. Vi vil lage en form for støysignal, et signal som ligner på lyden fra et stort fossefall. Lyden kan tenkes sammensatt av svært mange kilder som opptrer uavhengig av hverandre (lyden fra små og store vannklumper som treffer steiner og vannspeil i bunnen av fossen). Siden lyden blir til i en mengde ikke-relaterte prosesser, kaller vi lyden for “tilfeldig” eller “stokastisk”. Lyden har mange frekvenskomponenter som dekker et helt frekvensbånd.

Vi kunne også i denne delen av kapitlet tatt utgangspunkt i hørbar støy vi finner i naturen

eller andre typer bølger med liknede karakteristikk. Vi velger imidlertid å generere de tilfeldige signalene (kaotiske bølgene) numerisk. Da er det enkelt å endre på karakteristikkene.

Vi velger en fremgangsmåte der vi lager frekvensbildet til signalet slik vi ønsker det skal være, og så bruker vi en invers fouriertransformasjon for å generere signalet i tidsbildet.

Vi velger at frekvensspekteret (frekvensbildet) skal ha mange frekvensbidrag med en valgfri senterfrekvens og en gaussisk fordeling av frekvenskomponenten for nærliggende frekvenser. Bredden for frekvensfordelingen skal være valgfri. For å få stor variasjon hver gang vi genererer et signal, lar vi hver frekvenskomponent ha en vilkårlig verdi mellom null og den verdien gaussfordelingen angir. I tillegg lar vi fasen på hver frekvenskomponent være vilkårlig mellom 0 og 2π .

En Matlabfunksjon som genererer et slikt vilkårlig signal med gitt senterfrekvens og gitt fullverdibredde i frekvensfordelingen er som vist nedenfor.

```
function [xx] = HvitStoyGauss(Fs,N,fsenter,fullFbredde)

% Parametre: Fs : Samplingsfrekvens, N : Antall datapunkt
% fsenter, fullFbredde : Frekvensspekteret har en gaussisk fordeling,
% med senterfrekvens fsenter og full bredde (1/e) i
% frekvensspekteret lik fullFbredde..

fsigma = fullFbredde/2.0;
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));
ampl = rand(N,1);
ampl = ampl.*transpose(gauss);

faser = rand(N,1);
faser = faser*2*pi;
y = ampl.*(cos(faser) + i* sin(faser));

% Speiler nedre del rundt (Nhalv+1) for å få øvre del korrekt
Nhalv = round(N/2);
for k = 1:Nhalv -1
    y(N-k+1) = conj(y(k+1));
end;
y(Nhalv+1) = real(y(Nhalv+1));
y(1) = 0.0;

q = real(ifft(y)*200);
```

I figur 15.4 er det vist tre eksempler på vilkårlige signaler som er generert på denne måten, sammen med analyse av signalene med de samme tre analysemetodene som i stad. Det er valgt samme senterfrekvens (5000 Hz) for alle tre signalene, men tre ulike fullverdibredder (ned til 1/e av max), nemlig 50 Hz, 500 Hz og 5000 Hz. Det er brukt 2^{16} punkter og samplingsfrekvensen er 44100 Hz (samme som lyd på CD-er). Det framkommer en rekke interessante resultater.

Spraglet waveletdiagram

Det er interessant å se waveletdiagrammet for disse vilkårlige signalene. Allerede i figur 15.3 så vi at summen av flere signaler førte til at amplituden enkelte tidsperioder var stor, men liten i andre perioder. Vi fikk en ”oppklumping” av signalet i tid. I figur 15.3 var det imidlertid liten forskjell i frekvens.

Det ser vi bedre i figur 15.4. Oppklumpingen av signalet forekommer på en temmelig kaotisk måte både i frekvens og tid når bredden i frekvensfordelingen er stor nok. Det er opp til tusenvis av bidrag (frekvenskomponenter) til det endelige signalet (alle frekvenskomponenter innenfor frekvensfordelingen), og resultatet er altså et temmelig kaotisk spraglet mønster i waveletdiagrammet.

Dersom vi skalerte frekvenser og samplingsfrekvens i våre beregninger til synlig lys, ville den siste kolumnen i figur 15.4 kunne sammenlignes med elektromagnetiske bølger fra Sola. I så fall kunne det kanskje være fristende å få assosiasjoner til fotoner. Hver rød flekk i waveletdiagrammet kunne i så fall assosieres til et foton som kommer til en gitt tid og som har en viss frekvens. Men vi vet ut fra måten vi har generert dette signalet på at effekten skyldes en rett fram summasjon av mange vilkårlige bølger med en bredde i frekvensfordelingen. Dette er et fingeravtrykk av summen av mange samtidige uavhengige bølgebidrag, hvilket er helt naturlig siden lys som sendes ut fra enkelte deler av Sola skjer på en kaotisk måte og det er ikke noe korrelasjon mellom lys som kommer fra en del av soloverflaten med det som kommer fra andre deler.

Merk at bare et utsnitt av det totale waveletdiagrammet er gitt for signalene med størst bredde i frekvensfordelingen. Det viser at varigheten til hver røde flekk (periodene det finnes en betydelig amplitude for den frekvensen flekken tilsvarer) blir kortere når bredden i frekvensfordelingen øker. Det er vanskelig å anslå en slags gjennomsnittlig varighet for de røde flekkene, og resultatet er til dels også avhengig av valg av K-verdi for waveletanalysen. Likevel kan vi gi et (meget grovt) estimat av varigheten til flekkene slik:

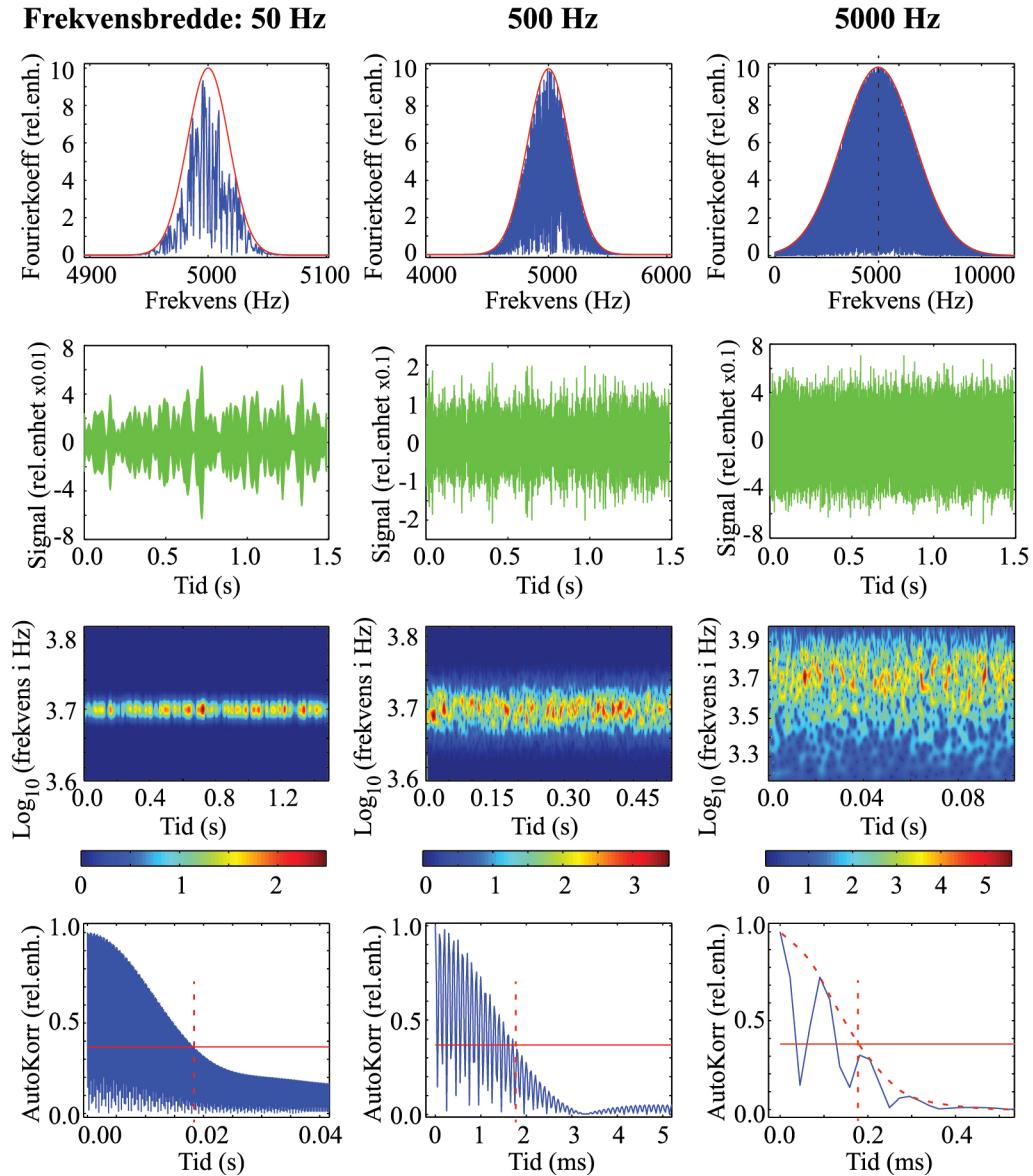
Frekvensbredde(Hz)	Varighet – av – rde – flekker(ms)
50	20 – 50
500	6 – 10
5000	1 – 2

Vi ser at varigheten til flekkene avtar når bredden i frekvensfordelingen øker. Vi merker oss videre at for den minste bredden i frekvensfordelingen er det ikke mulig å påvise at flere frekvenser opptrer samtidig, men når bredden i frekvensfordelingen øker, er det mange eksempler på at mer enn en frekvens kan være betydelig til stede samtidig.

Bredde i frekvensfordelingen vs koherenslengde

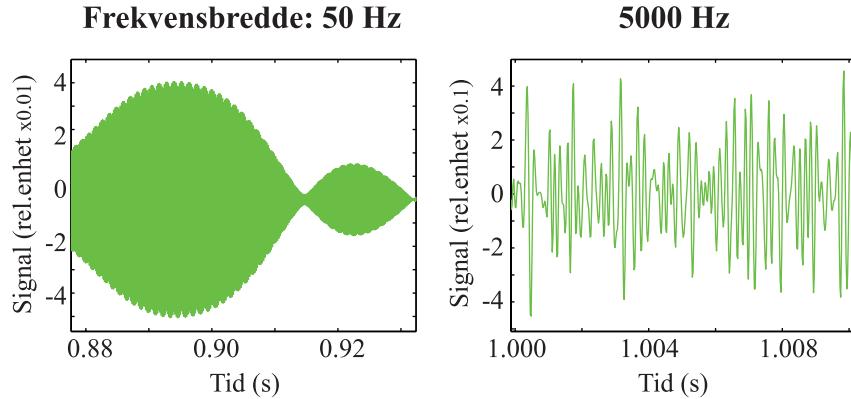
Vi finner følgende sammenheng mellom bredde i frekvensfordelingen og koherenslengden:

Frekvensbredde(Hz)	Koherenstid(ms)	Produkt – av – disse
50	18	0.9
500	1.8	0.9
5000	0.18	0.9



Figur 15.4: Tre eksempler på syntetiske kaotiske svingninger, med en bredde i frekvensspekteret på, fra venstre til høyre, 50, 500 og 5000 Hz. Senterfrekvensen er 5000 Hz. Det er også vist frekvensanalyse, waveletanalyse og den første delen av autokorrelasjonsfunksjonen. For waveletanalysen ble $K=120$, 36 og 12 brukt på analysene fra venstre til høyre (ca optimale verdier).

Det interessante er at dersom bredden i frekvensfordelingen er liten, tar det forholdsvis lang tid mellom forekomst av konstruktiv og destruktiv interferens (for tenkte delsignaler med litt forskjellig frekvens). Hver gang signalet får en økt amplitud, vil det ta en tid om lag lik den inverse av bredden i frekvensfordelingen før amplituden igjen avtar mot null.



Figur 15.5: Utsnitt fra signalet i tidsbildet for å vise at hver “boble” med relativt rent signal er mye større når frekvensbredden er liten enn når den er stor.

Figur 15.5 viser en detalj i tidsbildet for et signal generert med 50 Hz (venstre) og med 5000 Hz (høyre) bredde i frekvensfordelingen. Merk at for den minste frekvensbredden tar hver boble i tidsbildet i størrelsесorden 15-40 ms (mer enn 100 periodetider, hver svingning kommer ikke fram i plottet). Med hundre ganger større bredde i frekvensfordelingen er bredden på hver boble (i den grad det lar seg gjøre å definere slike) bare ca 0.3-0.5 ms (om lag to periodetider). Med velvilje kan vi si at en økning i bredden i frekvensfordelingen med en faktor hundre, førte til en reduksjon i varigheten i hver boble i tidsbildet med en faktor hundre. Dette er i tråd med sammenhengen mellom bredde i frekvensfordelingen og koherensstiden til signalet.

Resultatet henger sammen med det såkalte *Wiener-Khinchine teoremet* som sier at den fouriertransformerte av første ordens autokorrelasjonsfunksjon til en funksjon er lik frekvensspekteret til funksjonen. Vi går ikke inn i detaljer om denne siste relasjonen.

15.2.1 Sum av flere vilkårlige signaler, romlig koherens

Så langt når vi har omtalt vilkårlige signaler (kaotiske signaler), har vi bare studert hvert signal i seg selv. Ved beregning av koherensstiden har vi likevel på en måte sammenlignet et kaotisk signal med seg selv, litt forskjøvet i tid. Vi så at koherenslengden var meget liten når bredden på frekvensfordelingen var omrent like stor som sentrfrekvensen.

Vi skal nå studere *romlig* koherens, det vil si vi skal undersøke korrelasjonen mellom bølgen som passerer punkt A med bølgen som passerer punkt B i høyre del av figur 15.1.

Anta at det bare er én kilde til en bølge og at utstrekningen til kilden er mindre enn en bølgelengde. Anta videre at bølgen brer seg ut fra denne kilden med en tilnærmet kulesymmetri (i alle fall for den delen vi er interessert i). Anta videre at A og B er plassert like langt fra kilden. I så fall vil bølgen i A til ethvert tidspunkt være lik bølgen i B. En bølgetopp vil passere A og B samtidig. Vi kan si at A og B ligger på samme veldefinerte bølgefront, som er en del av en kuleflate med kilden i sentrum. I så fall er den romlige koherensen like lang som lengden til den delen av kuleflaten hvor denne relasjonen holder stikk.

Det betyr at vi kunne sette inn en dobbeltpalt vinkelrett på bølgens bevegelsesretning og kunne se interferens på en skjerm bak dobbeltpalten. Dersom koherenstiden er mer enn noen få periodetider, vil vi kunne få flere linjer i interferensbildet.

På tilsvarende måte kunne vi sette inn en kule og kunne påvise Aragos flekk (lys flekk) midt i skyggebildet av kula.

For lys er det imidlertid vanskelig å lage en lyskilde med en utstrekning mindre enn bølgelengden. Det er mulig i såkalte "quantum dots", men når en glødelampe sender ut lys, kan vi betrakte glødetråden som en mengde uavhengige lyskilder med kaotiske lyssignaler i hver kilde. Dersom vi nå studerer lysbølger som passerer to punkter A og B i "samme avstand" fra glødelampa, vil det ikke lenger være god korrelasjon mellom bølgene i de to punktene. Dette henger sammen med at det er en viss fordeling av avstander mellom de ulike delene av glødetråden og punktet A, og en *annen* fordeling av avstander mellom de samme delene av glødetråden og punktet B.

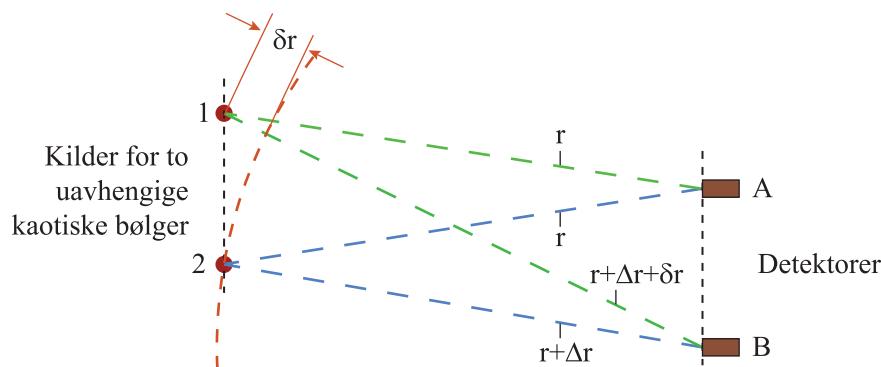
Samme tankegang kan brukes også for lys fra overflaten til en stjerne. Hanbury Brown og Twiss utviklet en elegant metode i 1954 og 1956 som kan brukes for å måle utstrekningen til en stjerne ved å benytte egenskaper til kaotiske bølger.

I figur 15.6 er prinsippet for Hanbury Brown og Twiss-effekten vist. Vi har for enkelhets skyld bare inkludert to uavhengige kilder med samme type kaotisk signal et stykke fra hverandre, og to detektorer A og B. Signalet inn til A består av summen av bølger fra kilde 1 og 2, og summeringen skjer uten at signalet fra kilde 1 er forskjøvet i forhold til signalet fra kilde 2.

Signalet inn til B er også en sum av signalet fra kildene 1 og 2, men denne gangen må signalet fra kilde 2 være forsinket i forhold til signalet fra kilde 1.

Signalene fra 1 og 2 er begge kaotiske og summen er også kaotisk. Flekkene i waveletdiagrammet vil ha kort varighet og være fordelt i tid og frekvens på en kaotisk måte (forutsatt stor bredde i frekvensspekteret).

Siden sumsigalet inn på A skjer uten tidsforskyving av delsignalet innbyrdes, mens sumsigalet inn på B skjer med en slik tidsforskyving, vil waveletdiagrammene for A og B kunne bli ganske forskjellig. Det betyr at det kan bli en dårlig korrelasjon mellom signalene inn til detektor A og B. Dersom vi derimot plasserte A og B på samme sted, ville korrelasjonen bli maksimal.



Figur 15.6: Prinsippskisse for å vise grunnlaget for den såkalte Hanbury Brown og Twiss-effekten. Se teksten for detaljer.

Selv om fordelingen av avstander i prinsippet alltid er forskjellig i punkt A og B dersom disse punktene ikke sammenfaller, er det ikke sikkert at forskjelligheten vil kunne påvises i praksis. Dersom forskjellene i fordelingen av avstander for A og B adskiller seg med godt mindre enn en bølgelengde, vil vi ikke kunne forvente å se noe forskjell i bølgene i punkt A og B. I så fall vil det være en høy grad av korrelasjon mellom bølgene i disse punktene. Gjør vi avstanden mellom A og B større, vil vi før eller senere få en forskjell på mer enn en bølgelengde i fordelinger av avstander fra ulike deler av lyskilden til punkt A og tilsvarende for B. I så fall vil vi kunne forvente at korrelasjonen mellom bølgene i A og B avtar.

Den største avstanden A og B kan ha som fortsatt gir en betydelig korrelasjon, sier vi er bølgens romlige koherenslengde på det stedet vi betrakter.

Sammenhengen er at dersom lyskilden, som består av uavhengige kaotiske partier, har en utstrekning gitt ved en vinkeldiameter θ fra det stedet vi betrakter lyskilden, og den gjennomsnittlige bølgelengden er λ , vil den romlige koherenslengden a være i størrelsесorden

$$a = \lambda/\theta$$

Hanbury Brown og Twiss brukte denne relasjonen i 1956 for å beregne størrelsen til stjernen Sirius som ligger 8.6 lysår vekk fra oss. Den romlige koherenslengden for lyset fra Sirius var om lag 8 meter her på Jorden. Vinkeldiametren ble anslått til 0.0068 busekunder som svarer til $3.3 \cdot 10^{-8}$ radianer.

Bruker vi denne relasjonen og betrakter en halogenpære (med 1 cm utstrekning på glødetråden) som lyskilde og betrakter lyset fra denne 100 m unna, vil vi kunne finne korrelasjon i punkter A og B som er opp til ca 6 mm fra hverandre. Med andre ord ville den romlige koherensen for denne lampa i 100 m avstand være om lag 6 mm.

[♣ ⇒ For det vilkårlige (kaotiske, stokastiske) lyset har vi hittil bare betraktet hvordan flekkene i et waveletdiagram endrer seg f.eks. når to kaotiske signaler adderes med og uten tidsforskyvning. Vi har da tatt utgangspunkt i signaler med den aktuelle frekvensen og bredden i frekvensfordelingen og beregnet korrelasjoner på utslag-vs-tid-nivå (første ordens korrelasjoner). Det fungerer utmerket når frekvensen er mindre enn noen få GHz, men vi har ingen detektorer som kan følge tidsvariasjonen i synlig lys ($6 \cdot 10^{14}$ ved 500 nm). Vi kan derfor ikke sample og beregne autokorrelasjonsfunksjonen for lysbølger.

Detektorene for lys er såkalte “square law detectors” som gir en respons proporsjonalt med *kvadratet* av amplituden til lyset som kommer inn. Detektorene kan ikke følge den momentane intensiteten som varierer like raskt som selve den underliggende sinusen, men gir en integrert intensitet over en betydelig lengre tid. Dette svarer til at detektorer for lys bare kan registrere tidsvariasjoner i intensitet i et frekvensområde som ligger under 1 GHz. Da Hanbury Brown og Twiss gjorde sitt berømte eksperiment på Sirius i 1956, var den tilgjengelige båndbredden på detektorer og forsterkere bare 38 MHz. Hvordan kunne de da følge de mye raskere forandringene som ligger i lyssignalene selv?

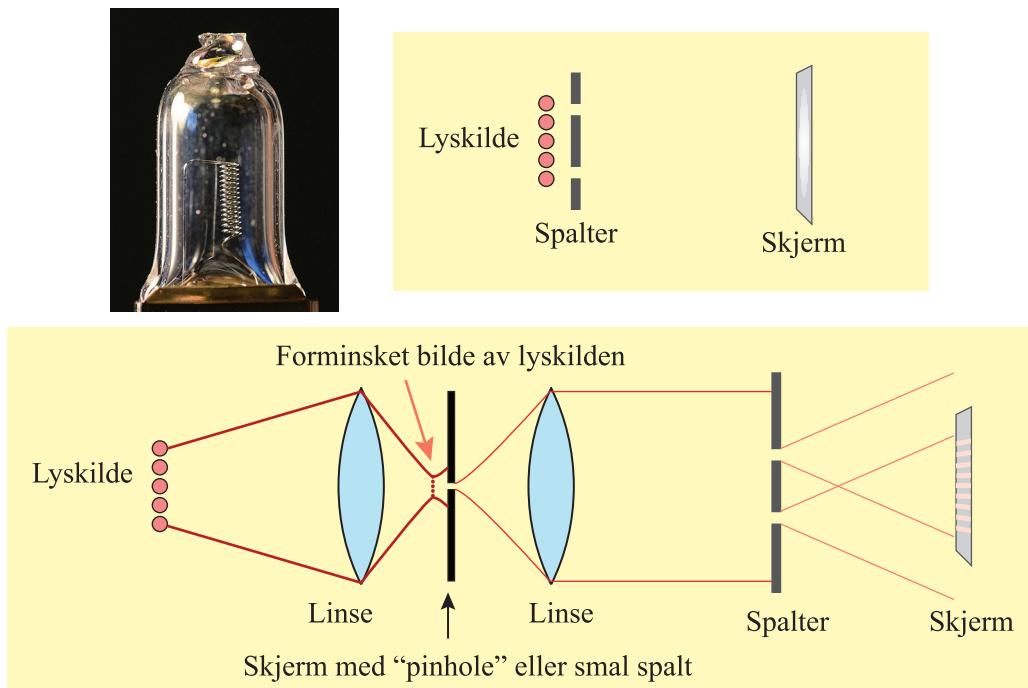
Det morsomme er at når vi summerer to frekvenser og kvadrerer summen, får vi følgende:

$$(\cos(\omega_1 t) + \cos(\omega_2 t))^2 = 1 + \frac{1}{2} \cos(2\omega_1 t) + \frac{1}{2} \cos(2\omega_2 t) + \cos((\omega_1 + \omega_2)t) + \cos((\omega_1 - \omega_2)t)$$

Foruten konstantleddet er frekvensen til tre av leddene om lag den dobbelte av de opprinnelige, og for deteksjon av lys er de helt utenfor mulighet for deteksjon. Leddet $\cos((\omega_1 - \omega_2)t)$ er imidlertid et slags “beat-frekvens-ledd”. For kontinuerlige frekvensfordelinger som vi har jobbet med her, vil dette leddet gi bidrag helt fra frekvensen null til en frekvens som svarer til bredden i frekvensfordelingen.

På grunn av dette “beat-frekvens-leddet” kunne Hanbury Brown og Twiss (og alle andre for den saks skyld) få transformert variasjonen som finnes i de opprinnelige waveletdiagrammene til et mye lavere frekvensområde.

Signalet som danner seg innenfor "beat-frekvens-området" kan også analyseres ved waveletanalyse, og vi får liknede kaotiske mønstre der også. Dette bildet gjør at korrelasjonen av signaler som angitt i figur 15.6 kan studeres også i de tilfellene at deteksjonen skjer med "square law detectors" i et helt annet frekvensområde enn de opprinnelige bølgene. ← ♠]



Figur 15.7: Vi får ikke interferensstriper når lys fra en halogenpære sendes direkte gjennom en dobbeltspalt. Dersom vi imidlertid først sender lyset fra pæra gjennom et "pinhole" med svært liten diameter, kan vi få tilstrekkelig stor romlig koherens til at interferensstriper kan påvises. For å få tilstrekkelig mengde lys gjennom hullet og nok lys til å se stripene på en skjerm eller på en fotografisk plate, er det en fordel å bruke konveks linser både før og etter hullet.

Romlig koherens i lyset fra Sola

Bruker vi samme relasjon for lyset fra Sola betraktet fra Jordas, finner vi at den romlige koherenslengden bare er om lag $60 \mu\text{m}$. Det betyr at det er umulig å bruke sollys direkte for dobbeltspaltekspesimenter og for å påvise Aragos flekk. Hva gjorde Young i 1801 og Arago ca 1820? Hemmeligheten ligger i å gjøre vinkelutstrekningen til lyskilden vi faktisk benytter liten nok. Det kan vi oppnå ved å sende lys gjennom et såkalt "pinhole". Vi kan lage et pinhole ved å stikke en meget tynn nål gjennom en aluminiumsfolie, eller vi kan kjøpe folier med ett veldefinert pinhole i en holder for en tusenlapp eller så. Vi kan velge mellom mange ulike diametere på hullet, og eksempelvis kan diameteren være $10 \mu\text{m}$. I så fall vil den romlige koherenslengden på lyset som slipper gjennom hullet være ca 5 cm etter at lyset har gått 1.0 m etter hullet (ved 550 nm).

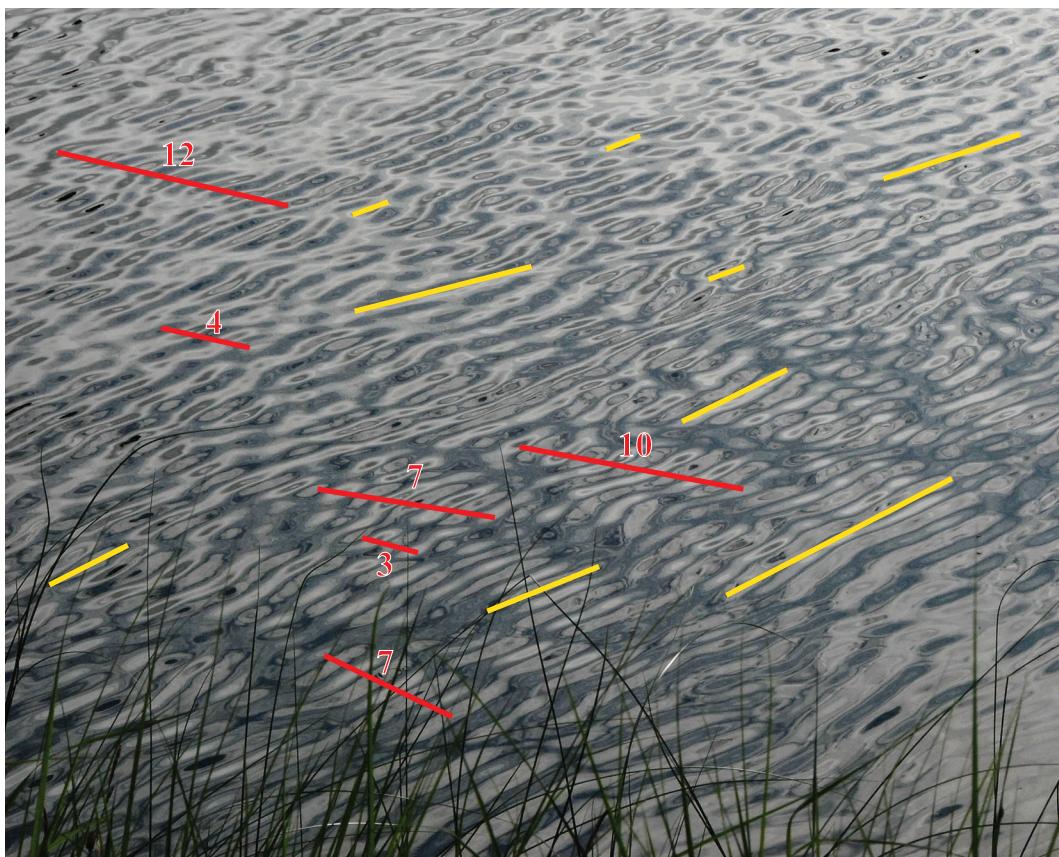
Det er imidlertid utrolig lite tverrsnitt i et hull med diameter $10 \mu\text{m}$. Dersom vi ønsker å gjøre et eksperiment med en glødelampe eller sola som lyskilde (ikke koherent lyskilde), kan vi øke intensiteten på lyset som slipper gjennom hullet ved å sende lyset gjennom en konveks linse og plassere hullet akkurat der bildet av sola (eller glødetråden) havner. Vi får også i

dette tilfellet bare benyttet en liten del av lyset fra kilden, men svært mye mer enn man ikke brukte en linse.

Det kan også være en fordel å bruke en linse etter hullet for å unngå at lyset sprer seg alt for mye ut. Denne siste lensen bør da plasseres slik at hullet er i lensens brennpunkt (se figur 15.7).

15.3 Anskueliggjøring av koherens

Det er ingen enkel sak å få en god forståelse av koherens. Vi velger derfor å ta med et fotografi fra overflatebølger på vann for å illustrere koherens på en annen måte enn hittil.



Figur 15.8: Overflatebølger på vann ved ett tidspunkt. I små flekker på overflaten har vi en temmelig “ren” bølge. Se teksten for videre omtale.

Innenfor små flekker på overflaten har vi en temmelig “ren” bølge (se figur 15.8). Innen disse flekkene er det mulig å forutsi relativt greit innbyrdes faseforhold i den retningen bølgen brer seg (røde streker). Innenfor flekkene er fasen og amplituden på bølgen omrent konstant i en retning normalt på den retningen bølgen beveger seg (gule streker). Flekkene har svært ulik størrelse. I bevegelsesretningen varierer flekkene mellom et par til tolv bølgelengder. Det betyr at den temporære (tidsmessige) koherenslengden er i størrelsесordne 5-7 bølgelengder, men det er slett ikke en nøyktig bestemt størrelse. I en retning normalt på bølgenvandringsretningen er de gule stripene i dette tilfellet gjennomsnittlig omrent like lange som gjennomsnittet for

de røde stripene. Det betyr at den spatielle (romlige) koherenslengden er omrent like lang som den temporære i dette tilfellet. Perspektiviske forhold gjør det imidlertid vanskelig å angi bredden på flekkene med nokså ensartete bølger.

Dersom vi betrakter bølger i tre dimensjoner, vil "flekkene" hvor det var temmelig veldefinerte bølger, bli erstattet med små volumer der det er temmelig veldefinerte (og tilnærmet plane) bølger.

Flekkene eller volumene med nokså veldefinerte bølger vil imidlertid endre seg i tid. Det gjør at kompleksiteten blir enda større. Det er ikke vanskelig å forstå at det er litt av en statistikk-utfordring å beskrive denne dynamiske situasjonen som vi svært ofte finner i praksis når bølger brer seg i rommet.

En liten detalj i figur 15.8 kan det være verdt å minne om. I ethvert punkt har vannoverflaten i et bestemt øyeblikk en nokså veldefinert høyde. Eller sagt på en annen måte: Vannoverflaten har ikke flere verdier samtidig! Vi har så lett for å tenke oss at det "finnes flere bølger samtidig", men på ett og samme sted har det lokale lufttrykket bare én verdi i et gitt øyeblikk for lydbølger i luft, og på ett og samme sted har det elektriske feltet bare én verdi og bare én retning i et øyeblikk for summen av samtlige elektromagnetiske bølger på dette stedet.

Dette er en egenskap det kan være vel verd å grunne litt over!

15.4 Måling av koherenslengde for lys

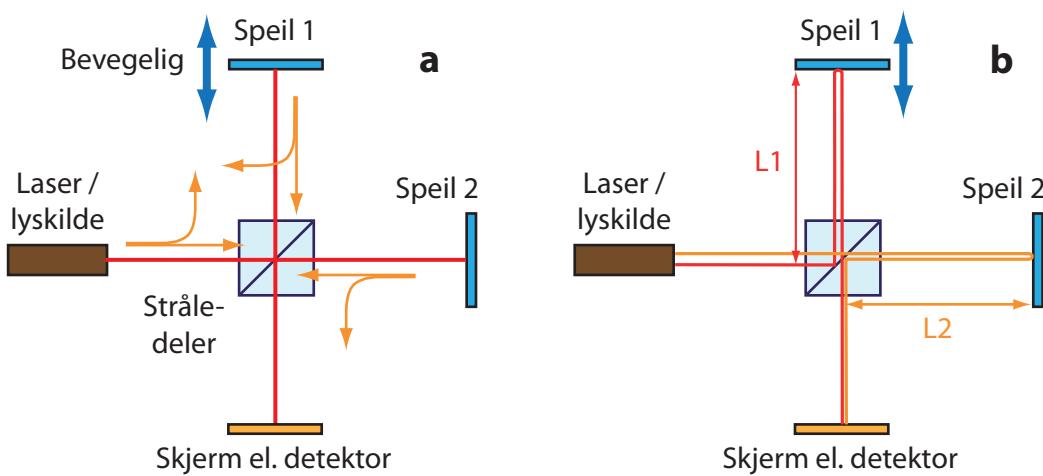
Synlig lys har en så høy frekvens at vi ikke klarer å registrere sinussvingningen av elektrisk felt etter som bølgen passerer. Vi kan da ikke bruke matematikken angitt ovenfor direkte.

Vi kan imidlertid gjennomføre en *analog* beregning av en størrelse som er nærmest beslektet med autokorrelasjonsfunksjonen. Det gjøres ved at en lysbunt splittes i to ved en såkalt "stråledeler" ("beam splitter"). De to delstrålene føres så sammen igjen, men først etter at den ene går en lengre vei enn den andre. Når strålene føres sammen, adderes elektrisk felt fra de to lysbuntene og magnetisk felt fra de to lysbuntene, og vi betrakter intensiteten av summen. Vi betrakter rett og slett:

$$\begin{aligned} G(\tau) &= 1/T \int_0^T (f(t) + f(t + \tau))^2 dt \\ &= 1/T \int_0^T (f^2(t) + 2f(t)f(t + \tau) + f^2(t + \tau)) dt \\ &= 1.0 + 2/T \int_0^T f(t)f(t + \tau) dt \end{aligned} \tag{15.7}$$

Her er $f(t)$ å betrakte som f.eks. elektrisk felt i strålen etter at den er delt i to og at amplituden er normalisert til 1.0.

Det betyr at vi rett og slett kan endre på gangveien til den ene dellysstrålen sammenlignet med den andre før de kombineres, og at vi da får ut autokorrelasjonsfunksjonen akkurat som i venstre del av figur 15.2, bortsett fra at hele kurven er forskjøvet med +1 slik at minimum ligger på null (intensiteten kan ikke være negativ). Figur 15.9 viser prinsippet i et såkalt Michelson interferometer som ofte brukes i slike målinger. Gangforskjellen for de to delstrålene er $\Delta L = 2L_1 - 2L_2$.



Figur 15.9: I et Michelson interferometer sendes en lysstråle inn til en stråledeler. Halvparten av strålen går til et fast speil og blir reflektert herfra, mens den andre halvparten går til et flyttbart speil. Halvparten av lyset som reflekteres fra speilene sendes til en skjerm eller detektor, hvor elektrisk felt fra de to bidragene adderes. I høyre del av figuren er lysveien for de to delstrålene markert skjematisk.

Lys fra termiske lyskilder, f.eks. glødelamper, kan ha en koherenslengde på bare noen få bølgelengder (det vil si bare noen få mikrometer). Lys med liten koherenslengde kalles "ikke-koherent". Lys fra en god laser kan ha en koherenslengde på opp til flere hundre meter. En laser som koster noen få tusen kroner har typisk en koherenslengde på noen centimetre (dvs i størrelsesorden 100 000 bølgelengder). Lys med lang koherenslengde kalles "koherent". Det er ingen skarp grense mellom ikke-koherent og koherent lys.

[♣ ⇒ Albert Abraham Michelson (1852-1931)

var en eminent eksperimentalfysiker. Han er kanskje mest kjent i forbindelse med Michelson-Morley-eksperimentet i 1887 som viste at lys ikke brer seg gjennom en ether. Michelson målte lyshastigheten med stor presisjon. Videre utviklet han stjerne-interferometre og klarte derved å måle diameteren på fjerne stjerner, og å måle avstanden mellom stjernepar ("binary stars" på engelsk). Han fikk i 1907 Nobelprisen i fysikk, og var den første amerikaner som fikk denne prisen. ← ♣]

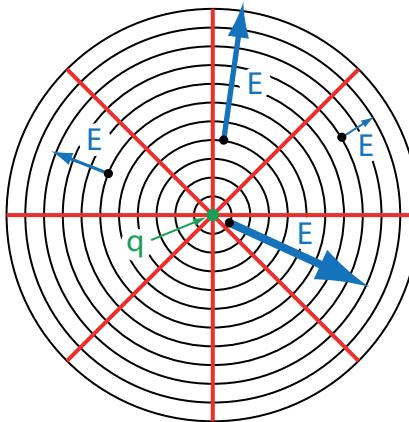
15.5 Stråling fra en elektrisk ladning

Koherens er knyttet opp mot mekanismene for hvordan bølgene genereres. Vi har tidligere såvidt diskutert mekanismer for å få bølger på en streng, lydbølger og overflatebølger på vann. For elektromagnetiske bølger har vi hittil bare konstatert at f.eks. plane bølger er løsninger av Maxwellts ligninger i fjernfeltsområdet i vakuum (i alle fall uten frie ladninger). Men hva er vanligvis selve *kilden* eller mekanismen bak generering av elektromagnetiske bølger? Vi skal bare såvidt berøre dette enorme feltet innen fysikk. Først skal vi se hvordan ladning i bevegelse kan gi bølger, og dernest skal vi se litt på noen av hovedtrekkene bak laseren.

Vi kan beregne hvordan vi kan lage elektromagnetiske radiobølger ved å sende en veksle-

strøm inn mot en antenn. I dette tilfellet har vi med frie ladninger og frie strømmer å gjøre, og Maxwells ligninger gir oss en ikke-homogen annen ordens partiell differensialligning for det elektriske feltet \vec{E} og en tilsvarende ligning for magnetfeltet \vec{H} . Beregninger av denne typen kan gjøres med Finite Element Metoder såvidt nevnt tidligere. Vi går ikke inn i detaljer her.

Vi velger her en “tegner og forteller”-fremstilling i stedet for en rigid matematisk behandling, men håper det vil være tilstrekkelig for å få fram et hovedtrekk.



Figur 15.10: En ladning i ro har sfæriske ekvipotensialflater rundt seg (sorte sirkler), og radielle elektriske feltlinjer (røde streker). Det elektriske feltet er tegnet inn i fire punkter i planet. Feltet er sterkt nær ladningen og avtar med økende avstand..

Figur 15.10 viser skjematiske at en ladning q i ro har elektriske feltlinjer som peker radielt utover (dersom q er positiv). Ekvipotensialflatene er kuleskall med sentrum i ladningen.

Beveger ladningen seg med konstant hastighet, vil ekvipotensialflatene ifølge relativitetsteorien bli “sammenklemt”, det vil si svakt diskosformet med minste akse i retningen bevegelsen skjer. I systemet hvor ladningen er i ro, er det bare et elektrisk felt. I et system hvor ladningen er i bevegelse, vil det være både et elektrisk og magnetisk felt. Når vi derimot snakker om å generere bølger, må vi trekke inn såkalte *retarderte potensialer*. En ordentlig behandling av dette temaet følger i senere kurs. Vi ser bare på noen overfladiske trekk her.

Vi bygger på en antakelse at *endringer* i elektrisk og magnetisk felt forflytter seg i rommet med lyshastigheten. Vi ser ikke en supernova når den skjer, men først etter at lyset har kommet seg den lange veien fra novaen til oss. Det vi ser i dag er supernovaen slik den var for akkurat tiden d/c siden, der d er avstanden mellom oss og supernovaen og c er lyshastigheten.

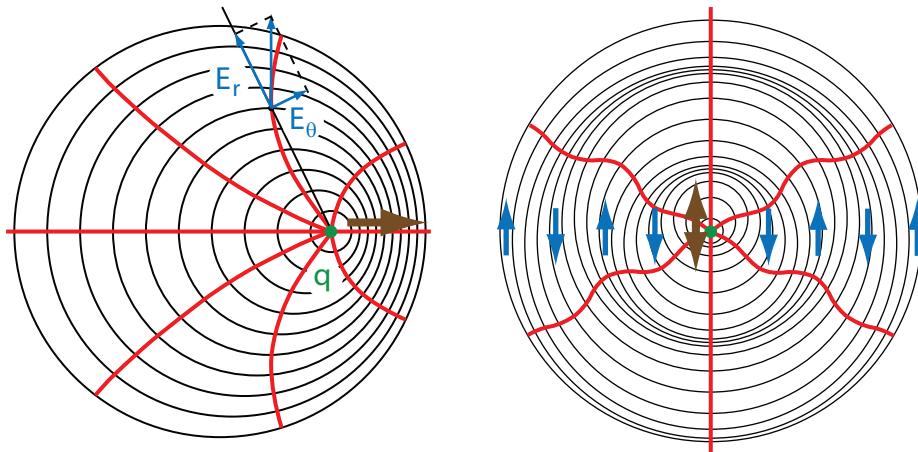
Slik er det også når vi flytter en ladning i rommet. Feltet et sted i rommet har da en fordeling som svarer til ladningens plassering på tidspunktet

$$t' = t - d/c$$

der t er tiden “nå” og d er avstanden fra ladningen til punktet vi betrakter feltet i, på tidspunktet t' .

Dersom vi da tegner opp ekvipotensialflater fra en ladning med konstant hastighet, vil flatene ha relative posisjoner som antydet i venstre del av figur 15.11. Effekten er sterkt overdrevet idet den aktuelle ladningen faktisk ville ha en hastighet over halve lyshastigheten slik figuren nå er tegnet.

Det elektriske feltet er normalt gitt som gradienten til det elektriske potensialet, og den hovedregelen beholder vi også når vi bruker retarderte potensialer. Da får vi elektrisk feltlinjer som er krummet, som vist i figuren.



Figur 15.11: Venstre del: En ladning i en konstant rettlinjet bevegelse har sfæriske ekvipotensialflater rundt seg, men bare i den forstand at ekvipotensialflaten i en viss avstand har sentrum i ladningens posisjon på tidspunktet d/c tidligere, der d er avstanden fra ekvipotensialflaten og ladningen på det tidligere tidspunktet. Forskyvningen av ekvipotensialflatene i forhold til hverandre fører til at de elektriske feltlinjene ikke lenger blir rent radielt rettet, men får en tangentiell komponent i tillegg. Feltlinjene i samme retning (og motsatt retning) som ladningen beveger seg, får ingen tangentiell komponent. Høyre del: En oscillerende ladning vil gi ekvipotensialflater som ligger litt forskjøvet i forhold til hverandre, som antydet. Det medfører at elektriske feltlinjer i tangentiell retning endrer seg som en bølge.

Anta at vi står i ro og en ladd partikkkel kommer farende forbi oss i konstant hastighet. Da vil vi oppleve et elektrisk (og magnetisk) felt på vårt sted som har en tidsutvikling der det elektriske feltet først har samme retning som ladningen beveger seg i (for positiv ladning), via et mye sterkere felt vinkelrett på denne retningen idet ladningen passerer, og ender som et svakt felt i motsatt retning av bevegelsen. Dette er en "puls" av elektrisk (og magnetisk) felt, og ikke en bølge i vanlig forstand.

En observatør som eventuelt fulgte med i ladningens konstante hastighet, vil beskrive det elektriske feltet som statisk, på en liknende måte som i vår figur 15.10. En slik situasjon kvalifiserer ikke for en utstråling av energi. I vårt eget referansesystem, hvor ladningen er i bevegelse, vil elektrisk felt bygges opp på ett sted i rommet, mens det skjer en helt ekvivalent nedbygging av felt et annet sted i rommet. Riktig nok vil området som har høyest feltenergi forflytte seg, på samme måte som ladningen, men denne forflytningen er av lokal karakter, og representerer ikke energi som fjerner seg fra området rundt ladningen.

For å få en bølge som brer seg ut over nærområdet til ladningen, må vi etterstrebe en situasjon lignende den vi hadde for en elektromagnetisk bølge i kapittel 9. Elektrisk (og magnetisk) felt må oscillere og ha en retning vinkelrett på bølgens bevegelsesretning. For å få til dette ved vår ladning i bevegelse, må vi ha en ladning som utsettes for en *akselerasjon*. Ladningen kan f.eks. oscillere fram og tilbake i rommet, gjerne i en harmonisk bevegelse. Det elektriskefeltet et stykke fra vil da oscillere som skissert i høyre del av figur 15.11. Denne tidsendringen i elektrisk felt vil ha både en radiell og en tangentiell komponent relativt til radiusvektor fra ladningen til punktet vi betrakter.

Komponenten i *radiell* retning (når vi er et stykke unna ladningen sammenlignet med utslaget i ladningens oscillasjon) vil (nesten) ikke endre seg med tiden. Denne komponenten vil derfor (nesten) ikke gi opphav til noe bølge som brer seg utover.

Komponenten *vinkelrett på den radielle retningen* vil derimot svinge (nesten) som en sinus med tiden. Denne komponenten vil kunne gi opphav til en elektromagnetisk bølge som brer seg ut i rommet.

Krumningen på det elektriskefeltet øker med hastigheten til ladningen mens den oscillerer. Den tidsderiverte av denne igjen bestemmer hvor stor $\frac{\partial E}{\partial t}$ blir. Disse to faktorene tilsammen medfører at utstrålt energi blir proporsjonal med frekvensen til svingningen i annen potens. Vi sier derfor ofte at utstrålingen er proporsjonal med *akselerasjonen* til ladningen.

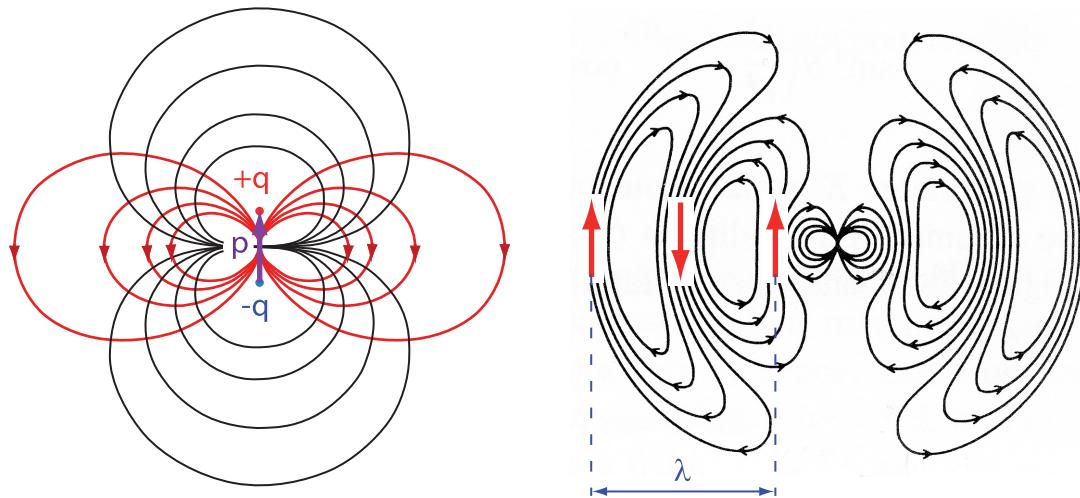
♠ ⇒ [Noen side-bemerkninger:

Det er nærliggende å tenke seg at elektrisk felt fra en ladning ”stråler” utover hele tiden. Det ligger liksom i kortene siden vi opererer med retarderte potensialer der vi tenker oss at feltet på et sted skyldes ladningen der den var for en stund siden. Et konstant utstrålende elektrisk felt ville imidlertid lett kunne stride mot energibevaring m.m. Heldigvis er det ikke nødvendig å tenke i de baner. Det er *endringer* i elektrisk og magnetisk felt som forplanter seg med lyshastigheten. Før en bestemt endring har bredd seg ut og nådd fram til et gitt sted, er det feltfordelingen som har opphav i forhold *før* endringen som gjelder. Det elektriskefeltet fra en ladning i ro er i likevekt med seg selv. Det er en løsning av Maxwells ligninger, og det skjer ingen endringer i felt og ingen transport av energi. Først så snart bevegelser og spesielt akselerasjon forekommer, blir ting annerledes.] ⇐ ♠

15.5.1 Dipolstråling

En alternativ måte å generere elektromagnetiske bølger på er å bruke en elektrisk (eller magnetisk) dipol som varierer i tid. Dette er en meget effektiv måte å lage bølger på. Det kan vi skjonne ved å betrakte elektrisk feltfordeling fra en permanent elektrisk dipol (se venstre del av figur 15.12). Det elektriskefeltet er rettet vinkelrett på den radielle retningen i planet normalt på dipolens retning.

Dersom vi endrer polariteten til dipolen på en harmonisk måte, får vi et elektrisk felt i dette ekvatorialplanet som vil variere akkurat slik vi ønsker det for å generere en elektromagnetisk bølge som kan bre seg ut i rommet (elektrisk felt vinkelrett på bevegelsesretningen). I retningen dipolen peker (og motsatt retning) er det elektriskefeltet et stykke fra dipolen omrent lik null, og det har ingen komponent på tvers av den radielle retningen. I disse to retningene blir det praktisk talt ikke utsendt noen bølger.



Figur 15.12: Venstre del: En statisk elektrisk dipol består av to identiske ladninger, men med motsatt fortegn, plassert litt fra hverandre. Det er tegnet inn såvel ekvipotensialflater (sorte, med buk oppover og nedover) som elektriske feltlinjer (røde, med buk utover mot sidene). Dipolens fysiske utstrekning er kraftig overdrevet i forhold til feltlinjemønsteret. Høyre del: En oscillatorende elektrisk dipol vil skape et elektrisk felt i rommet rundt som indikert. Feltmønsteret beveger seg utover med lyshastigheten. (Høyre del er en videreføring av en figur i P.Lorrain, D.R.Corsen, F.Lorrain: *Electromagnetic fields and waves*, 3rd ed.)

I høyre del av figur 15.12 er det vist et diagram over elektrisk feltfordeling nær en dipolantenne ved et gitt tidspunkt. Det elektriskefeltet er kraftigst der feltlinjene ligger tettest. Hele mønsteret beveger seg utover med lyshastigheten, og nye looper danner seg nær antennen, to ganger for hver periode (retning på feltet skifter retning i de to systemene av looper som dannes hver periode i dipolvariasjonen). En animasjon av forløpet (og mye annen informasjon) er tilgjengelig på Wikipedia under oppslagsordene “dipole radiation”.

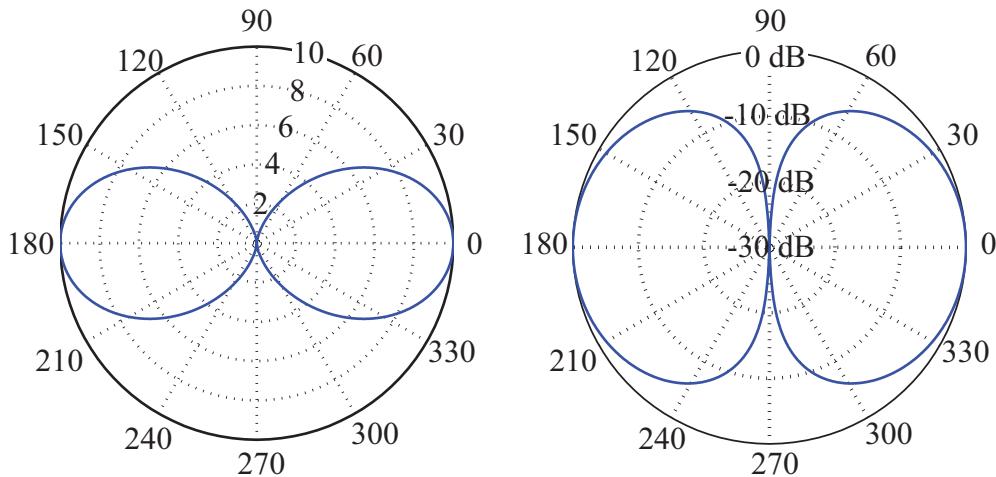
♠ ⇒ [En kommentar:

Høyre del av figurene 15.11 og 15.12 har en viss relasjon til hverandre, men er likevel forskjellige. Forsök å påpeke forskjeller og likheter.] ← ♠

Det er vanlig å tegne *retningsdiagram* for antenner. Et retningsdiagram angir relativ tidsmidlet intensitet på de utsendte bølgene for ulike retninger i rommet, i et såkalt “polar-diagram”. Figur 15.13 viser retningsdiagrammet i et vertikalplan som går gjennom en enkel vertikal dipolantenne. Et retningsdiagram kan gis med en lineær skalering i radiell retning (venstre del av figuren), men mest vanlig brukes det en logaritmisk skala i radiell retning (høyre del av figuren).

♠ ⇒ [En kommentar:

Anta at vi lager et polardiagram med logaritmisk skalering i radiell retning. I så fall må vi velge hvilken intensitet som skal svare til sentrum i diagrammet og skalere verdiene slik at denne intensiteten skaleres til 1.0 ($\log_{10}(1.0) = 0.0$). I figur 15.13 har vi valgt at en intensitet lik 1/1000 av maksimal verdi skal svare til null utslag. Intensiteter mindre enn dette ville bli negative i en logaritmisk skala og ville dukke opp på motsatt side av diagrammet. For å unngå misforståelser fjerner vi negative verdier før plotting.



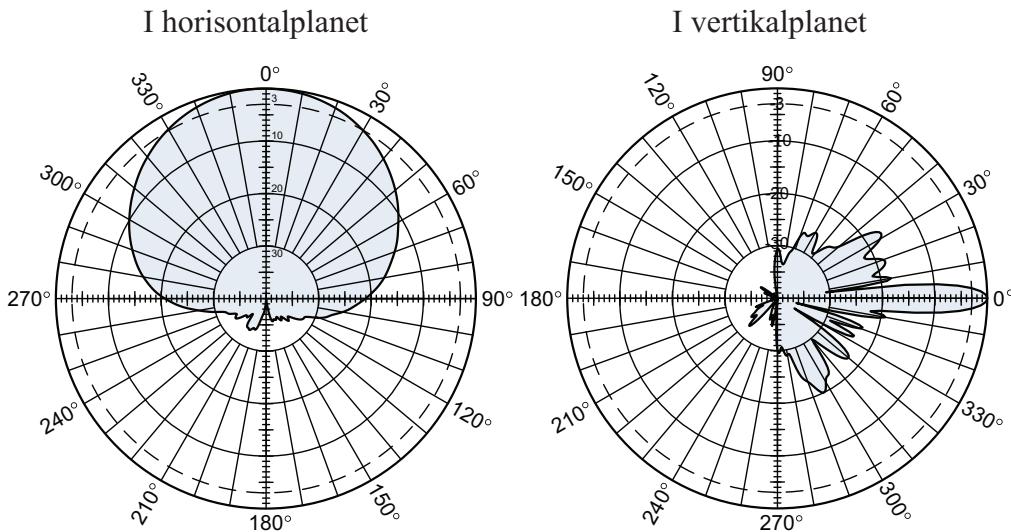
Figur 15.13: Retningsdiagram for en enkel vertikal dipolantenne. Diagrammet gjelder bare i en avstand fra antennen som er stor i forhold til antennens lengde. Det er brukt lineær skala for intensiteten i radiell retning i venstre del av figuren og en logaritmisk skalering til høyre. Intensitetene er alle relative til den maksimale verdien. Diagrammet leses slik (venstre del): Intensiteten ved 0 grader er satt til "10". Da er intensiteten ved 30 grader ca 7.3 og ved 60 grader ca 2.8.)

Matlabprogrammet som ble brukt for å lage figur 15.13 var som følger:]

```
function antennediagram3
N = 1024;
theta = linspace(-pi/2.0,3.0*pi/2.0,N); % Vinkler
costheta = cos(theta);
intensitet = costheta.*costheta;
intensitet = log10(intensitet*1000.0);
for i = 1:N
    if(intensitet(i)<0)
        intensitet(i)=0;
    end;
end;
polar(theta,intensitet);
```

En dipolantenne kan på sett og vis sees på som en enkeltspalt med svært liten bredde. Utstrålingen blir identisk i alle retninger vinkelrett på dipolens retning. Setter vi imidlertid inn to dipoler ved siden av hverandre, og mater begge antennene med identisk signal, vil strålingsdiagrammet se ut omtrent som for en dobbeltpalt. Ved å sette mange identiske antenner etter hverandre på en rekke, får vi et strålingsdiagram omtrent som for en enkeltspalt. Ved å sette inn reflektorer og direktorer, kan vi påvirke strålingsdiagrammet ytterligere, og et eksempel er gitt i figur 15.14. Det er et antennediagram for en antennen som brukes mye i basestasjoner for mobiltelefoni her i landet. Merk at diagrammet er radielt med desibelmål i radiell retning.

Det er langt på vei samme type tenkning som ligger bak utstrålingsdiagrammene for antenner som for lysintensitetsfordeling ved en og flere spalter osv. Det bærende element i beregningene er interferens mellom tilstrekkelig koherente bølger. Da kan vi bruke tenkningen



Figur 15.14: Retningsdiagrammer for en vanlig basestasjonsantenne for GSM 900 mobiltelefoni her i landet (Kathrein 80010621). Det ene diagrammet gir vinkelfordeling i vertikal retning, det andre for horisontal retning.

om forskjeller i veilengde for å addere ulike bidrag med korrekt innbyrdes fase.

15.6 Lasere

Lasere er i dag en av de viktigste lyskildene i naturvitenskap, og lasere brukes også i teknologi så som CD og DVD-spillere og i laserskrivere. Lasere brukes ved kutting av metaller og andre materialer, og i medisinen f.eks. ved omforming av hornhinna på øynene våre og andre operasjoner. Ja, til og med min tannlege har gått over til å bruke lasere ved “boring” i tennene. Noen bilfabrikanter bruker nå lasere som frontlykter. Bruksområdet for lasere er imponerende, og øker stadig!

Ordet laser er en forkortelse for Light Amplification by Stimulated Emission of Radiation. Theodore Maiman ved Hughes Research Laboratories klarte å lage verdens første laser (se figur 15.15). Dette skjedde 16. mai 1960. Laseren feiret derfor sitt 50-årsjubileum i 2010. Det er imidlertid svært mange fysikere som har vært involvert i utviklingen og utnyttelsen av laseren, og det var Charles H Townes som i 1964 fikk Nobelprisen i fysikk “for the development of laser principles”. Også andre Nobelpriser i fysikk er temmelig nært knyttet opp mot laseren på en eller annen måte. Det er derfor naturlig at vi kommer litt inn på prinsippene bak en laser. Det er imidlertid bare hovedprinsipper som blir berørt, og også her brukes en “tegner og forteller”-tilnærming.

En laser er basert på såkalt *stimulert emisjon*. Einstein hadde allerede i 1917 vist at vi kan ha stimulert emisjon fra f.eks. atomer. Med det mener vi at vi ikke behøver å eksitere et atom og vente til det finner det for godt å sende ut lys idet atomet går tilbake til grunntilstanden. Vi kan ved å tilføre litt lys til atomet faktisk *trigge/stimulere* atomet til å gå tilbake til grunntilstanden. Laseren er basert på at et medium lar seg stimulere til emisjon og at lyset som emisjonen fører til i sin tid er med på å stimulere utsendelse av enda mere lys fra mediet. På denne måten får vi en positiv tilbakekobling i prosessen som gjør at den nærmest går av



Credit: HRL Laboratories

<http://davidkirkpatrick.wordpress.com/category/etal/page/2/>

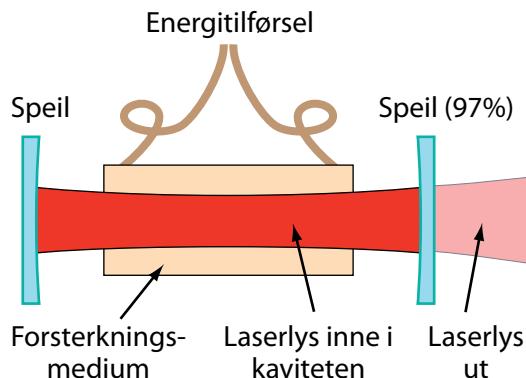
Figur 15.15: Foto av den første laseren. Et blitz-rør omkranser en rubinstav. Rubinstaven er belagt med et nærliggende 100 % reflekterende speil i ene enden og ca 95 % reflekterende speil i andre enden. Laseren ga fra seg pulset koherent lys. Fotografiet ble frigitt fra Hughes Research Laboratories i forbindelse med 50-årsjubileet i 2010.

seg selv. Energien som går til spille på ulikt vis må imidlertid kompenseres for, så det trengs en ytre energikilde dersom prosessen skal vedvare.

Figur 15.16 viser hovedingrediensene i en tradisjonell laser. Den inneholder et forsterkningsmedium som kan tilføres energi fra en ytre energikilde. Mediet ligger i en optisk kavitet ("hulrom" eller "boks") begrenset av to speil. Lys som dannes i mediet vil i utgangspunktet spre seg ut i alle retninger, men lys som treffer speilet har en tendens til å bli reflektert tilbake gjennom mediet, treffer speilet på andre siden, og går så fram og tilbake og setter opp en stående bølge av elektriske og magnetiske felt i kaviteten. Lyset som reflekteres fram og tilbake kan virke som stimuli for å få mediet til å gi fra seg enda mere lys.

I vanlig lys kommer lyset fra mange atomer eller molekyler som opptrer temmelig uavhengig av hverandre. Da får vi en tilstand som svarer til våre reelle sangere og våre kaotiske bølger nevnt tidligere i dette kapitlet. Lyset fra lyskilden øker omtrent proporsjonalt med antall atomer/molekyler som sender ut lys, og lyset kommer i "alle" retninger.

Atomene/molekylene som sender ut lys i en laser, derimot, befinner seg i kaviteten mellom de to speilene. Det elektromagnetiske feltet vil øke gradvis og mer og mer lys blir dannet på grunn av emisjonen. Når vi starter laseren er det spontan emisjon i atomene som dominerer, men etter hvert blir det mer og mer stimulert emisjon. En viss tid er det fortsatt mindre lys som dannes enn det som forsvinner i form av ulike typer tap. Før eller senere overgår stimulert lys tap av lys, og da løper systemet på en måte løpsk. Effekten av lyset i kaviteten stiger med flere størrelsesordener i en bestemt retning, og linjebredden til emisjonen blir flere størrelsesordener mindre enn den var fra først av. Den magiske grensen kalles på engelsk "lasing threshold", grensen for at laseren skal lase.



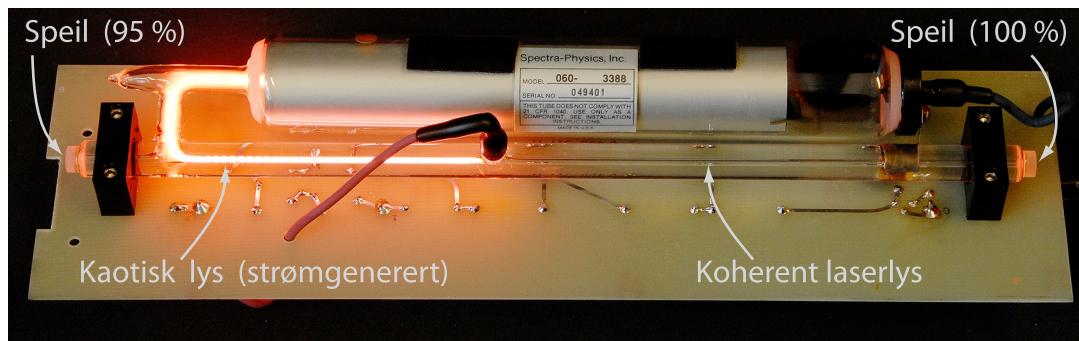
Figur 15.16: Hovedbestanddelene i en tradisjonell laser (prinsipielt).

Grunnen til at hele prosessen skal gå som beskrevet er at det kraftige elektromagnetiske feltet i kavitten sørger for at alt lyset som kommer fra de ulike atomene har temmelig lik fase som de stående bølgene mellom speilene. Da vil amplitudene av elektrisk og magnetisk felt adderes direkte, og intensiteten på lyset innenfor selve strålen blir kvadratet av antall atomer/molekyler som sender ut lys.

Siden lysbølgene i en laser befinner seg i en "kavitet" (hulrom med speil i begge ender, se figur 15.17), vil lyset som nevnt danne stående bølger. Da vil frekvensen bli meget presis, på en lignende måte som lyden fra en gitarstreng som er festet i begge ender er temmelig presis. Dersom avstanden mellom speilene er 30 cm for en HeNe-laser med bølgelengde ca 633 nm, vil det være om lag 473940 bølgelengder mellom speilene. Selv linjebredden på den energiovergangen vi bruker i neon-gassen (kolliderer med andre molekyler) er såpass bred at det da iblant kan forefinnes flere samtidige bølgelengder i kavitten. Med 473940 bølgelengder mellom speilene vil bølgelengden være 632.9915 nm, men med én bølgelengde mer eller midt enn dette, vil bølgelengden være 632.9902 og 632.9929 nm henholdsvis. Vi snakker om "moder" for laserlyset. Ved mekanisk oppvarming av laserkavitten, vil avstanden mellom speilene endre seg litt. I så fall vil bølgelengdene også endre seg, og vi får såkalt "mode hopping".

I enkelte sammenhenger lages det lasere som spiller aktivt på de ulike modene laseren kan operere i. Vi kan da oppnå mange bølgelengder som ligger med tilnærmet samme bølgelengdeavstand fra hverandre, og fenomenet kalles "frekvens-kam". Theodor Hänsch fikk i 2005 Nobelprisen fordi han hadde laget en "frequency comb synthesizer" som gjorde det mulig for første gang å måle oscillasjonene i lys med ekstrem presisjon. Metoden danner basis for våre mest moderne atomur.

Selv om frekvensen til laseren skulle endre seg litt etter som tiden går, blant annet fordi temperaturfluktasjoner gjør at lengden på kavitten endrer seg litt med tiden, vil ikke dette ødelegge for hovedprosessen. Alt lyset som kommer fra ulike atomer/molekyler inne i laseren vil fortsatt være i fase fordi feltet i kavitten trnger nytt lys som sendes ut til å være i fase med hovedfeltet til enhver tid. Av den grunn vil aldri et av atomene bidra med lys som er i motfase til lys fra et annet atom. Lyset blir koherent. Dette er en idealisert beskrivelse, men virkeligheten ligger ikke så langt fra det ideelle i gode lasere.



Figur 15.17: Et fotografi som viser innmaten til en vanlig laboratorielaser av HeNe-typen. En fortynnet blanding av helium og neon befinner seg i en kavitet mellom to speil. En del av kaviteten er felles med et mer eller mindre standard lysrør a la det vi finner i farget reklamebelysning natterstid ("neon-reklame"). Elektrisk strøm gjennom dette lysrøret gir kaotisk, ikke-koherent lys som sendes ut i alle retninger. Energi tappes fra de atomene som er eksitert av den elektriske strømmen og kollisjonene den medfører. Denne energien brukes for å bygge opp en kraftig, koherent lysstråle mellom de to speilene. Vi ser ikke denne strålen fra siden, fordi lyset i kaviteten er nær perfekt rettet langs aksen mellom speilene, og da slipper ikke stort ut til sidene.

Laserlyset slipper ut av kaviteten ved at ett av speilene i enden av kaviteten ikke reflekterer 100 %, men bare kanskje 95 - 98 %. Det lyset som da slipper ut er ganske annerledes enn lys fra f.eks. en glødelampe.

Dersom vi sammenligner fasen til den elektromagnetiske bølgen rundt omkring i et plant tverrsnitt normalt på strålen, vil fasen overalt i planet være nesten identisk. Vi omtaler dette som romlig (spatsiell) koherens. Høy grad av romlig koherens vil si at bølgefronten er meget plan. Dette er ulikt lys som stammer fra mange atomer som ikke har noe faserelasjon med hverandre, f.eks. lys fra en glødelampe. Slikt lys har liten romlig koherens, hvilket betyr at bølgefronten bukler mye på seg.

Den stimulerte emisjonen i kaviteten er temmelig stabil, og frekvensen er så veldefinert hele tiden, at vi kan forutsi fasen i laserlys-strålen mange bølgelengder framover. Denne form for koherens kalles temporær eller tidsmessig koherens.

Siden bølgefronten er svært godt definert på tvers av strålen samtidig som vi kan forutsi fasen til laserlyset lange avstander langs selve strålen, er en laser en ekstremt mye bedre lyskilde ved interferens og diffraksjons-eksperimenter sammenlignet med såkalt termisk lys ("ikke-koherent" lys). Det betyr også at en laserstråle vil holde seg samlet som en meget vel avgrenset stråle der diffraksjon holdes på et minimum. Lyset i en laserstråle er noe av det nærmeste vi kan komme til en matematisk idealisert bølgebeskrivelse i praksis. Laserlys kalles derfor iblant for "klassisk lys", men en slik betegnelse forvirrer mer enn den er til nytte.

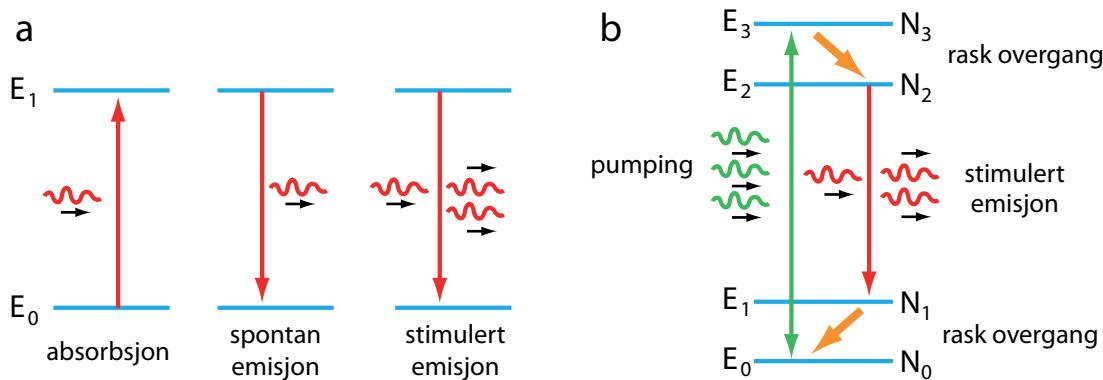
15.6.1 Populasjonsinvertering

Når vi skal forklare lasere, kommer vi ikke helt utenom en detalj kalt populasjonsinvertering. Vi skal ikke gå i detalj, siden dette temaet ikke er så viktig i vårt sammenheng. Her er likevel en kort gjennomgang.

Et atom kan befinner seg i flere ulike energitilstander. Vi tegner ofte energitilstandene skjematisk som i venstre del av figur 15.18. Grunntilstanden markeres gjerne med E_0 og

første eksitere tilstand E_1 . Et atom kan eksiteres fra grunntilstanden til en eksiteret tilstand bl.a. ved å plassere det i et elektromagnetisk felt med frekvensen $\nu = (E_1 - E_0)/\hbar$ der \hbar er Plancks konstant. Et atom i eksiteret tilstand kan falle tilbake til grunntilstanden helt av seg selv (kalles spontan emisjon). Vi kan også stimulere overgangen med et elektromagnetisk felt med samme frekvens som angitt for absorpsjon.

Ved absorpsjon stjenes lys fra en lysstråle for å eksitere atomet, mens ved stimulert emisjon frigjøres lys fra atomet. Det er samme sannsynlighet for den ene overgangen som den andre per atom, forutsatt at det er i den aktuelle utgangstilstanden. Skal vi få frigjort mer lys enn vi putter inn (slik det kreves i en laser), må det være flere atomer i den eksiterete tilstanden enn i tilstanden atomet faller tilbake til ved emisjon. Populasjonen av energinivåene følger vanligvis Boltzmannstatistikk. Da er det flere atomer i en lav energitilstand enn i en høyere. Det vil si at det vanligvis er umulig å danne en laser ut fra atomer i termisk likevekt.



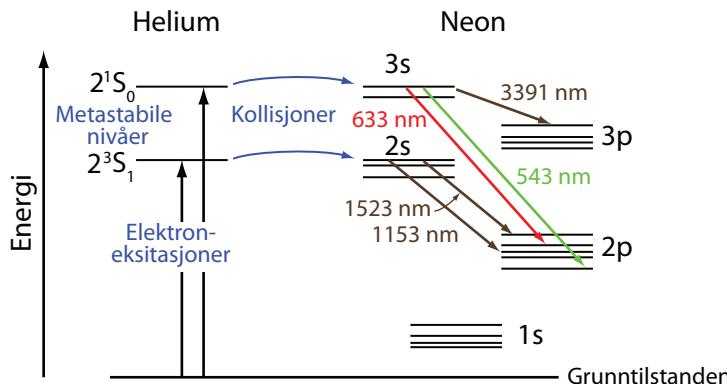
Figur 15.18: Venstre del: To energitilstander i et atom, og skjematiske overganger mellom disse (meget forenklet). Høyre del: Populasjonsinvertering kan oppnås ved pumping mellom andre energinivåer enn hvor laserlyset skapes. Se tekst for detaljer.

I høyre del av figur 15.18 er det vist en måte å få til høyere populasjon i en energitilstand enn i en lavere. Prinsippet brukes i neodymium YAG lasere og er basert på fire energinivåer. Atomene eksiteres ved hjelp av kraftig lys fra en eller annen lyskilde, fra grunntilstanden til fjerde energinivå (E_3). Atomet går da spontant raskt over til energitilstanden E_2 , men holder seg her. Det er også en rask spontan overgang fra E_1 ned til grunntilstanden. Overgangen fra E_2 til E_1 er imidlertid ikke raskt, og etter en del pumping, blir det flere atomer i E_2 enn i E_1 . Vi har fått en populasjonsinversjon!

Sender vi nå et (svakt) lys med frekvensen $\nu = (E_2 - E_1)/\hbar$, vil vi få mer utsendt lys enn absorbert lys fra atomene, og det ligger an til at vi kan danne en laser. Intensiteten på laseren vil likevel være begrenset av hvor raskt vi kan pumpe atomer fra grunntilstanden til E_4 når en laserovergang har funnet sted.

En av de mest vanlige laserne i laboratoriesammenheng er helium-neon laseren. Her er mer kompliserte energinivåer involvert. En skisse (til orientering) er vist i figur 15.19. I dette tilfellet eksiteres heliumatomer ved at det sendes en elektrisk strøm gjennom gassblandinga av helium og neon. Elektroner med betydelig hastighet står for eksitasjonen. Helium har to eksiterete nivåer som er "metastabile" slik at helium kan være i disse tilstandene ganske lenge

før de faller ned til lavere energier. Dersom et slikt eksitert heliumatom kolliderer med et neonatom, vil den eksitere energien kunne overføres fra helium til neon. Neonatomet kan så deeksiteres videre bl.a. gjennom en overgang som gir lys ved 632.8 nm. Det er dette røde lyset vi kjenner igjen fra en HeNe-laser.



Figur 15.19: Energitilstander som er involvert i en vanlig HeNe-laser.

I dag er det mange ulike måter å lage en laser på. Folk flest har vel opp til flere lasere i sitt hjem, i og med at CD og DVD-spillere bruker lasere. I tillegg har mange også en laserpeker. I alle disse eksemplene benyttes halvleder-laserdioder. Lyset fra slike laserdioder er kontinuerlig i tid. Det finnes også lasere som bare gir fra seg til dels meget korte lyspulser. Pulslengden kan være helt ned i det såkalte femtosekund-området, ja til og med ned i det såkalte attosekund-området (bruker denne betegnelsen for pulser litt kortere enn 1 fs). Bølgelengden er ikke vel definert for slike korte laserpulser!

Relevans for oss?

[♣] ⇒ Da vi gikk gjennom generering av elektromagnetiske bølger ved hjelp av en oscillerende ladning eller oscillerende dipol, baserte vi oss på Maxwells ligninger. Prosessen ble beskrevet som kontinuerlige funksjoner, og vi fikk en elektromagnetisk bølge som varte ved så lenge oscilleringen foregikk.

Når vi forklaarte laseren brukte vi energinivåer og hopp fra en energitilstand til en annen. Et slikt bilde bygger på kvantefysikk, men egentlig bare en kvantefysikk som er basert på energiegentilstander der vi bruker pertubasjonsteori for å se på sannsynligheter for overganger. Hvordan skal slike energidiagrammer som angitt i figur 15.18 oppfattes? Når en overgang først foregår, foregår den da øyeblikkelig, eller tar overgangen en del tid? Dette spørsmålet er det vanskelig å få et godt svar på!

Vi tegner gjerne inn "fotoner" som små bølgepakker i slike diagrammer, hvilket innebærer at vi antyder at det kommer ut en liten bølge når et foton frigjøres fra et atom. Men hvor lang er så denne bølgen? Kan vi ha bølger som kommer ut som har en fasehukommelse (koherenslengde) som svarer til flere hundre tusen bølgelengder, men som selv har bare nesten ingen utstrekning?

Og hvordan kan det ha seg at elektroner ved lavere frekvenser gir en kontinuerlig, vedvarende bølge i Maxwells formalisme, mens elektromagnetiske bølger ved lys-frekvensene plutselig ikke kan beskrives som vedvarende bølger (men som foton-partikler)?

Det er en kobling mellom kvantemekanikk og det klassiske bildet av oscillerende elektriske dipoler: Det er nettopp dipolmomenter i atomer, beregnet kvantemekanisk etter en operator som har en klar klassisk analogi, at vi kan regne ut sannsynligheter for at et atom skal sende ut lys.

Til tross for at fysikere i dag ofte omtaler lys som "fotoner" som har en litt uklar partikkelnatur, kan de aller fleste fenomener hvor lys er involvert forklares ut fra bølgemodellen for lys. Det er svært få fenomener hvor vi må bruke en partikkelmønster. Men hva menes med bølger og hva menes med partikler når alt kommer til alt? Kan det hende bølge-partikkkel-dualismen og de tilsvynelatende paradoksene som følger av en slik oppfatning, kan forsvinne dersom vi forsøker å bli litt mer presise i vår beskrivelse?

Det er nå vel 100 år siden forrige gang fysikere skiftet fra en grunntanke til en annen grunntanke om lys. Kanskje det er på tide at det skjer et nytt skifte? ↛ ♠]

15.7 Læringsmål

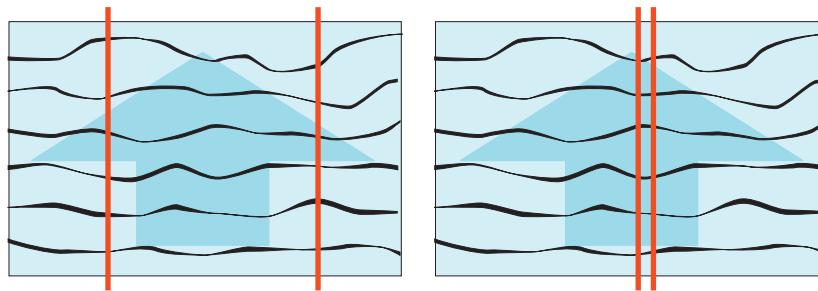
Etter å ha jobbet deg gjennom dette kapitlet bør du kunne:

- Gjøre rede for hva som skiller en reell bølge fra en idealisert enkel matematisk beskrivelse av en bølge.
- Gjøre rede for hva som menes med tidmessig (temporær) koherenstid og koherenslengde, og hvordan de kan bestemmes.
- Gjøre rede for hva som menes med romlig koherens.
- Kunne forklare hvorfor koherens spiller en rolle når vi skal gjøre interferenseksperimenter.
- Forklare kvalitativt at linjebredde i et frekvensspekter har sammenheng med koherenslengder.
- Forklare hvordan vi kan måle koherenslengder med et Michelson interferometer.
- Forklare kvalitativt at en ladning i oscillerende bevegelse fører til at det sendes ut elektromagneiske bølger.
- Angi en kvalitativ sammenheng mellom en sammensatt radiofrekvens-antenne og diffraksjon av lys fra to eller flere spalter i en skjerm.
- Forklare hvorfor en laser i utgangspunktet får en (temporær og longitudinal) koherenslengde som er langt større enn termisk lys fra f.eks. en glødelampe.
- Forklare kvalitativt hvorfor populasjonsinversjon er viktig for å få en laser til å fungere.

15.8 Oppgaver

Forståelses- / diskusjonsspørsmål

1. Forsök å forklare hva som ligger bak at signalet i midtre del av figur 15.3 varierer så mye i amplitude selv om signalene vi startet ut med hadde mye jevnere amplitude.
2. Dersom du har sunget i kor har du sikkert lagt merke til at lydstyrken når flere synger sammen ikke øker så mye som vi kanskje skulle tro. En solist med god stemme høres slett ikke så verst selv i vekselsang med et kor på kanskje 30 personer. Hvorfor blir ikke den ene stemmen totalt overdøvet av de 30?
3. Vil en sanger som synger en "død" tone forventes å ha en større eller mindre koherenstid for stemmen enn en sanger som har betydelig "vibrato" i sangen sin?
4. Angi fordeler og ulemper ved at folk som synger i et kor har relativt korte koherenstider for stemmen (lyden).
5. I figur 15.20 er det markert linjer som skal svare til posisjonene til spaltene i en dobbeltpalt i forhold til bølger som passerer forbi. Vil du tro at du vil kunne påvise interferensstriper for tilfellet til venstre og/eller tilfellet til høyre? Forklar hva som ville skje i de to tilfellene.



Figur 15.20: Markering av posisjon til de to spaltene i en dobbeltpalt i forhold til bølgemønsteret. Avstanden mellom spaltene er mye større i eksemplet til venstre enn til høyre.

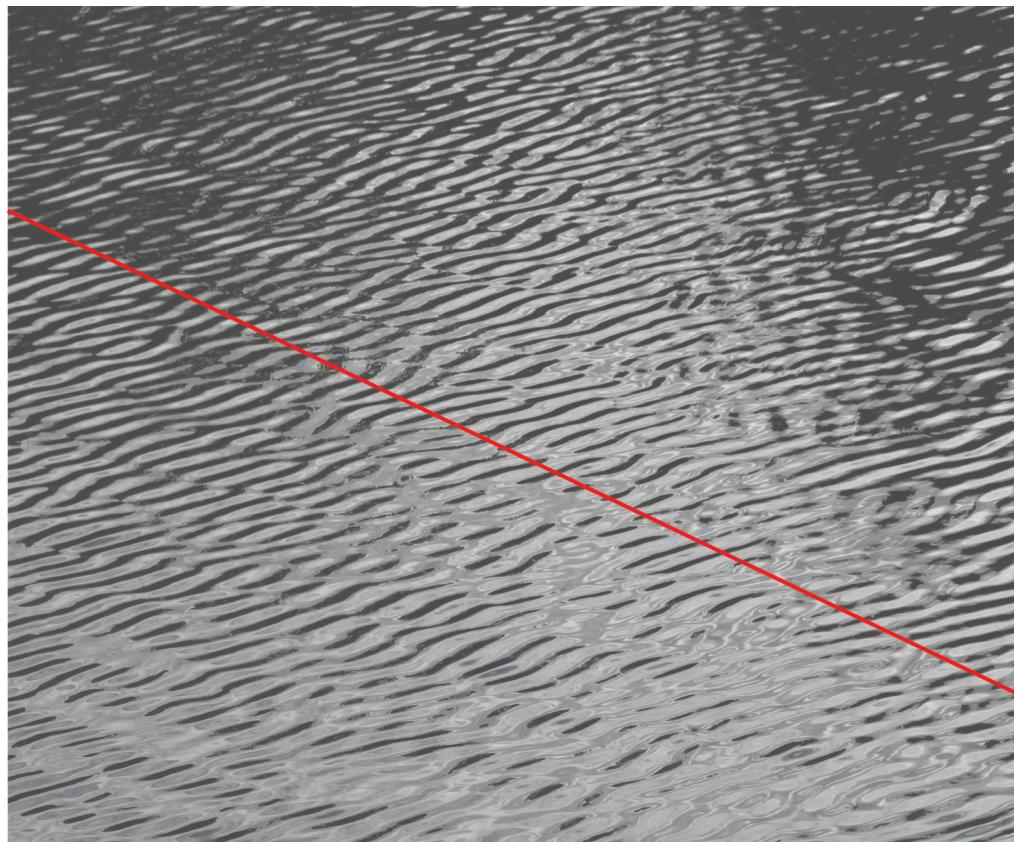
6. Noen tror at når vi snakker om koherente og ikke-koherente bølger, er det snakk om to vel avgrensede typer bølger. I virkeligheten er det en kontinuerlig overgang fra "ikke-koherent" til "koherent". Forklar.
7. Det er fordeler og ulemper knyttet til både koherent lys og ikke-koherent lys. I hvilke sammenhenger ville du foretrekke det ene og når ville du foretrekke det andre? Begrunn som vanlig svarene.
8. Hva må til for å vise minst fem interferensstriper etter å ha sendt lys gjennom en dobbeltpalt, når lyset vi bruker kommer fra Sola?
9. Forsök å beskrive med egne ord hva vi mener med koherenslengde og koherenstid for en bølge. Hva er forskjellen mellom romlig koherens (spatiell koherens) og tidsmessig koherens (temporær koherens)?
10. Hva er hovedideen bak metoden til Hanbury Brown og Twiss forenklet angitt i figur 15.6?
11. Hva er hovedtanken bak "retardert potensial"?
12. Hva er viktigst for å skape en elektromagnetisk bølge fra en oscillerende ladning eller oscillerende dipol, enten den radielle komponenten av det elektriske feltet eller den

tangentielle komponenten? Forklar!

13. Forklar hvordan populasjonsinversjon kan oppnås i et fire-eneginivå-skjema.
14. Antenner er ofte utformet som elektriske dipoler som antydet i figur 15.12. I blant brukes imidlertid magnetiske dipoler som antenner (en enkel sirkulær strømsløyfe). Skissér magnetfeltet rundt en magnetisk dipol og bruk denne som utgangspunkt for å argumentere for at antennen kan forventes å være ganske effektiv.

Regneoppgaver

15. Bruk programkoden tidligere i kapitlet for å generere et kaotisk signal med senterfrekvens 5000 Hz og bredde i frekvensfordelingen på 3000 Hz. Lag en funksjon som kan beregne autokorrelasjonsfunksjonen. Plott denne. Er koherenstiden du kommer fram til omtrent som forventet? Kommenter resultatene.
16. Bruk programkoden tidligere i kapitlet for å generere et kaotisk signal med senterfrekvens 5000 Hz og bredde i frekvensfordelingen på 3000 Hz. Foreta en fourier- og waveletanalyse av signalet. Beregn så kvadratet av det opprinnelige signalet (elementvis kvadrering). Foreta en fourier- og waveletanalyse av det kvadrerte signalet (pass på å ta med et stort nok frekvensområde i waveletanalysen). Kommenter resultatene.
17. Bestem koherenstiden til egen stemme. Nærmere bestemt går oppgaven ut på følgende:
 - a) Lag et dataprogram hvor du kan digitalisere lyd, beregne autokorrelasjonsfunksjonen og plotte en utvalgt del. Bruk plottet til å anslå omtrentlig koherenstid for signalet. Dersom du har laget et program for digitalisering av lyd ved arbeidet med kapittel 5, kan du vinne mye tid på å utnytte dette også her.
 - b) Fortell spesielt hvordan du valgte å utnytte datastrengen du fikk ved digitaliseringen i analysen. Nærmere bestem: Hvordan valgte du å la indeksene i og j i ligning (15.5) løpe i forhold til den totale datastrengen?
 - c) Bestem omtrentlig koherenstiden til egen stemme når du synger “iiiii” med så jevn stemme du klarer. Gjennomfør dette for 2 - 3 ulike tonehøyder. Synes koherenstiden å endre seg mye med tonehøyden?
 - d) Foreta en waveletanalyse av signalet. Kommenter resultatet.
 - e) Digitaliser en annen lyd og bestem koherenstiden også for denne. (Forslag til lyd: Egen stemme, samme tonehøyde som du har brukt i punkt c, men at du nå synger “ooooooooo” i stedet for “iiiii”. Alternativt: Lyd fra et piano, gitar eller et annet musikkinstrument.) Finner du noe interessante forskjeller eller likheter sammenlignet med det du fant i punkt c?
18. Forsøk i figur 15.21 å markere områder hvor bølgene er relativt veldefinerte. Hvor store er disse områdene omtrentlig? Og hvor stor del av hele vannoverflaten har du tatt hensyn til i markeringen av disse områdene? Angi lengder i “omtrentlige bølgelengder” som mål. [Det er som hjelp tegnet inn en linje som angir omtrentlig retningen bølgene beveger seg i. For å få et korrekt bilde av koherenslengder må egentlig hele overflaten tas med i den statistiske behandlingen.]
19. a) I figur 15.14 er det gitt et antennediagram for en mye brukt basestasjonsantenne. I et konkret tilfelle står en slik antennepal på en mast 22 m over bakken. Bestem intensiteten på bakkenivå i en avstand 30 m fra masten (målt langs bakken) i forhold til intensiteten 500 m fra antennen i den retningen i horisontalplanet hvor intensiteten er størst. [Husk at radiell retning i antennediagrammet angir relativ intensitet i antall dB for bølger som



Figur 15.21: Overflatebølger på vann ved ett tidspunkt.

går ut i ulike retninger fra antennen. I høyre del av figuren svarer origo til at intensiteten er 40 dB lavere enn i maksimalretningen (0 grader).]

- b) Gjennomfør samme beregning dersom det i stedet var brukt en enkel dipolantenne.
 - c) Er det gunstig at basestasjonantennen har den intensitetsprofilen den har, eller ville det vært en fordel om det ble brukt en enkel dipolantenne i stedet?
20. a) Bruk "båndbreddeteoremet" (den klassiske ekvivalenten til Heisenbergs uskarphetsrelasjon) for å bestemme bredden i frekvensspekteret for en laserpuls som bruker 40 fs (femtosekunder) å passere et punkt i rommet (bilde fra laseren er vist på første side av dette kapitlet). Bølgelengden som svarer til midten i frekvensspekteret er 810 nm. [Pulsene har en nær gaussisk omhyllingskurve i både tid og i frekvensspekteret.]
b) Hvor mange bølgelengder er det i én puls?
c) Hvor "lang" er lyspulsen (målt i samme retning som lyset går i)? [Det blir fort rot dersom du blander relativitetsteori inn i slike beregninger, så vi anbefaler at du holder deg til en ikke-relativistisk beskrivelse.]
d) Gjenta de samme beregningene som ovenfor for en lyspuls med bredde på 7.7 fs. [Dette er en laser som for tiden brukes i München i deres forsøk på å lage attosekund-laser. Interesserte henvises til et oppslag 7. mai 2015 i Photonics.com med overskrift: "Laser Design Brings Attosecond Spectroscopy Closer" (<http://www.photonics.com/Article.aspx?AID=57412> nedlastet 10. mai 2015.)]

Det er interessant å se hva forskere bak dette arbeidet bl.a. skriver:

“This field of ultrafast physics focuses on phenomena such as electron motions in molecules and atoms, which can take place on attosecond timescales. The ability to generate attosecond laser pulses would effectively permit electron motions to be “photographed.””

Som du skjønner foregår det spennende utvikling i fysikk i dag!

21. Kan du foreslå en måte å lage et interferometer for lyd som i funksjon tilsvarer et Michelson interferometer for lys? (Skal kunne brukes for å måle koherenslengder for lyd fra f.eks. ulike musikkinstrumenter.)
22. Bestem koherenslengden for en laser ved å bruke et Michelson interferometer.
23. Søk på internett for å danne deg et bilde av status for lasere i røntgenområdet. Hvor langt er utviklingen kommet? Hvilke anvendelser har en røntgenlaser?