Sampling og rekonstruksjon

Øyvind Hammer
Norsk nettverk for Teknologi, Akustikk og Musikk

Oslo, 16. mars 1994

1 Diskretisering

Lyd er en analog signaltype, som må omformes til digital form for å kunne behandles av en digital datamaskin. Denne omformingen kalles *sampling*, og utføres av et apparat kalt en *analog/digital-omformer* (A/D-omformer).

Et analogt signal er *tidskontinuert*, det vil si et det kan endre seg kontinuerlig over tid og er definert til ethvert tidspunkt. Det er også *amplitydekontinuert*, dvs. at momentanutslaget kan anta enhver verdi innenfor et intervall som begrenses av systemets dynamiske område.

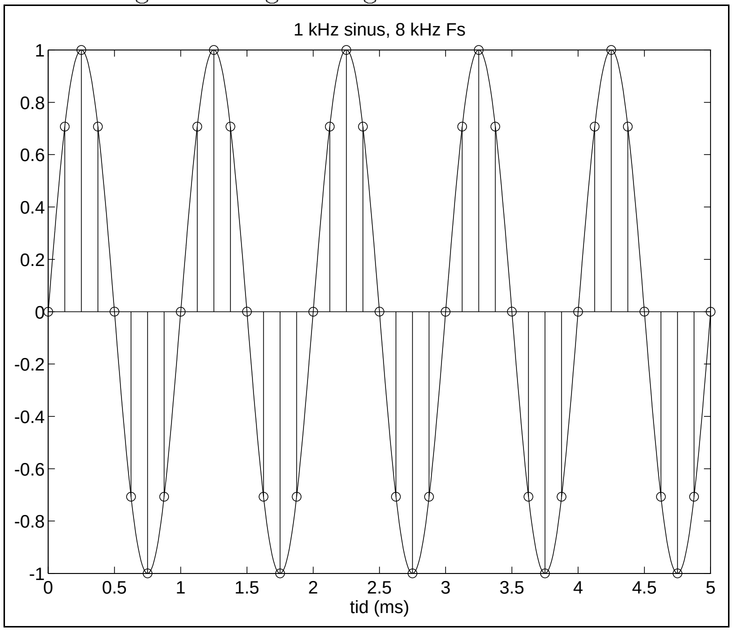
For å kunne behandles av en digital datamaskin med endelig hukommelse, ordbredde og regnehastighet, må signalet bringes over i *diskret form*. Dette vil si at signalet måles på et endelig antall tidspunkter (signalet gjøres tidsdiskret) og med et endelig antall mulige måleverdier (signalet gjøres amplitydediskret). Denne prosessen medfører en forringelse av signalet, som vi ønsker å minimalisere.

Omformingen fra et tidskontinuert til et tidsdiskret signal er en følge av at vi måler (sampler) signalet på et endelig antall tidspunkter. Hvilke verdier signalet antar mellom disse målingene kan vi ikke vite, og signalet er derfor ikke definert mellom samplingstidspunktene. I de systemer vi skal omtale, er tidsrommene mellom målingene konstante, slik at vi kan snakke om en konstant målefrekvens. Denne frekvensen kalles *samplingfrekvensen*. Hvis vi måler signalverdien 10 000 ganger i sekundet, sier vi at samplingfrekvensen er 10 kHz. Et sentralt spørsmål når vi skal sample et signal, er hvor høy samplingfrekvens vi bør bruke. Åpenbart ønsker vi å benytte en så lav samplingfrekvens som mulig, uten at signalet forringes for mye. Derved begrenser vi datamengden, og dermed kravene til datamaskinens hukommelse og regnehastighet.

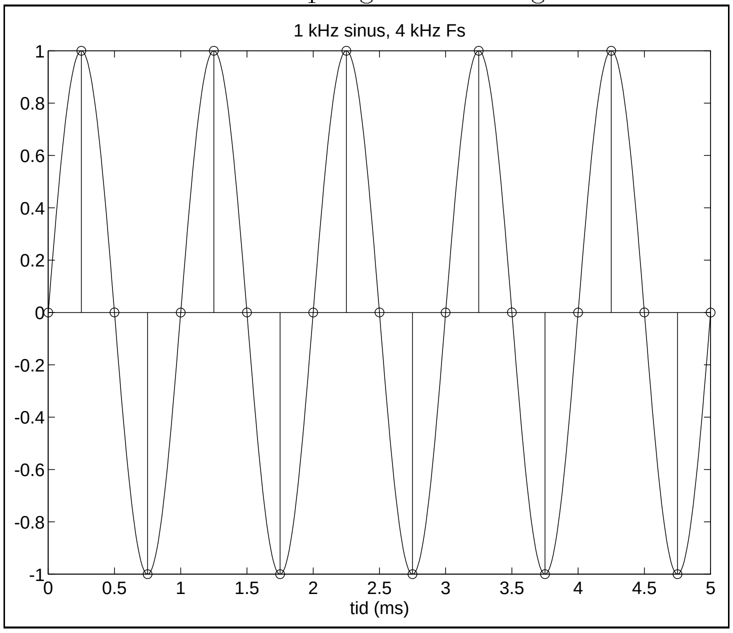
2 Samplingsteoremet, aliasering

En grunnleggende regel i samplingsteorien er at samplingfrekvensen må være mer enn dobbelt så høy som den høyeste frekvenskomponenten vi ønsker å kunne registrere. Sagt omvendt: Den høyeste frekvenskomponenten vi kan registrere er mindre enn halve samplingfrekvensen. Dette er en del av *Shannon og Nyquist's samplingsteorem*. Halvparten av samplingfrekvensen kalles *Nyquist-frekvensen*. Vi skal ikke teoretisere for mye rundt samplingsteoremet, men illustrere det ved et eksempel. Dette eksemplet vil også demonstrere det fryktede *aliaserings*-fenomenet.

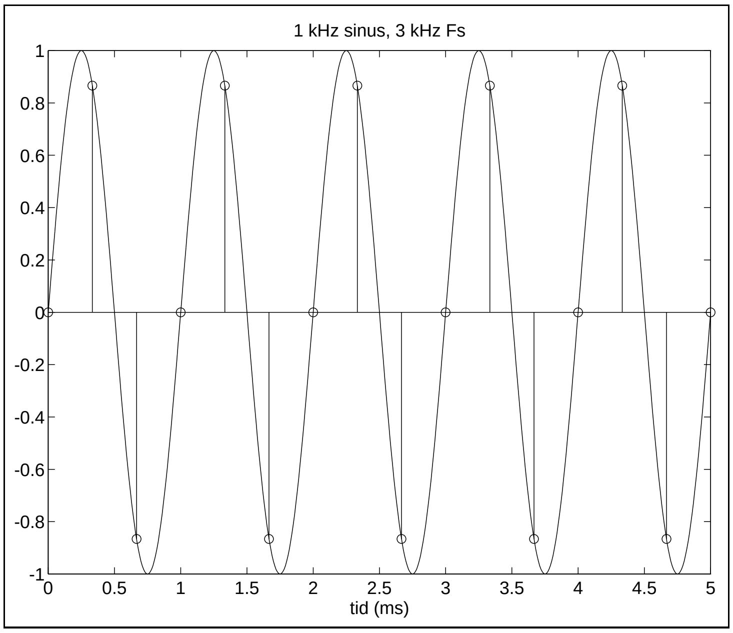
Figuren under viser et analogt, sinusformet signal, la oss si med frekvens 1 kHz (1000 Hz). Vi ønsker å sample dette signalet, dvs. måle momentanverdien med jevne mellomrom. Figuren viser signalet samplet med 8 kHz samplingfrekvens. De målte verdiene vises ved hjelp av stolper. Vi ser at signalet er godt registrert.



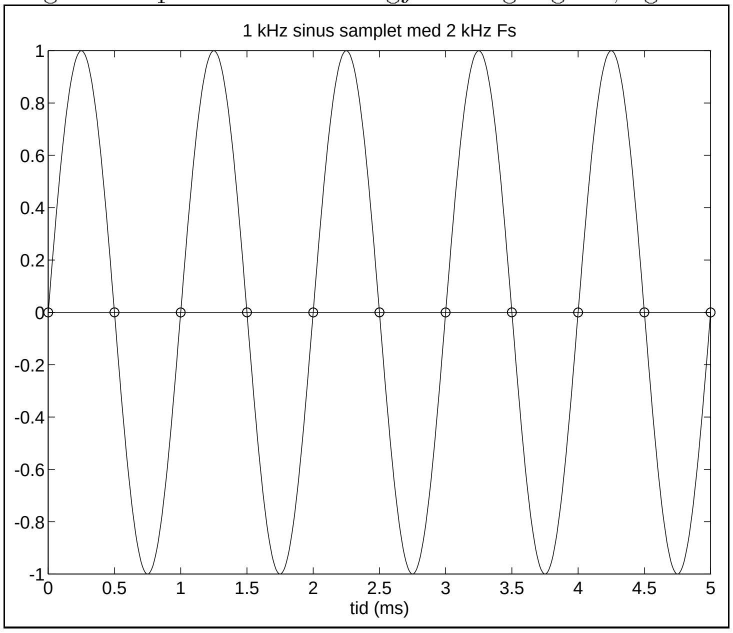
I neste figur går vi ned til 4 kHz samplingfrekvens. Signalet er fremdeles registrert.



Her er sampling med 3 kHz, signalet er fremdeles registrert:

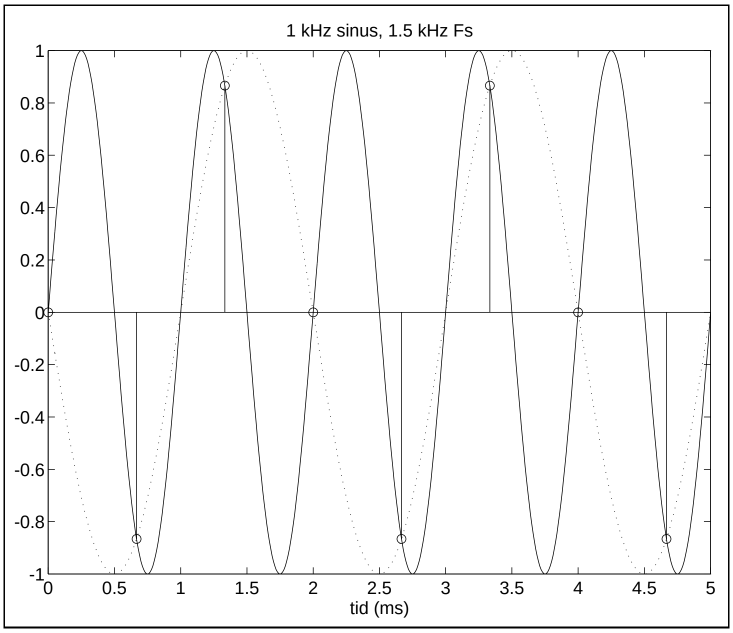


Så går vi helt ned til 2 kHz samplingfrekvens. Vi ser at det går dårlig. I dette spesielle tilfellet er vi så uheldige å sample akkurat i nullgjennomgangene, og vi får ikke ut noen ting.



En samplingfrekvens på 3 kHz er altså tilstrekkelig for å registrere en frekvenskomponent på 1 kHz, mens 2 kHz blir for lite. Dette stemmer med samplingsteoremet.

La oss så forsøke å senke samplingfrekvensen ytterligere. I figuren under sampler vi med 1500 Hz. Det samplede signalet beskriver nå en sinustone på bare 500 Hz (stiplet linje).



Signalet på 1000 Hz er altså blitt feilregistrert som et signal på 500 Hz. Med andre ord blir signaler som ligger over halve samplingfrekvensen foldet (speilet) nedover i spekteret, og danner falske frekvenser. Dette er et alvorlig problem i digital lyd, som kalles *aliasering*. Vi tar det en gang til:

Aliasering er et fenomen som kan oppstå i en samplingsprosess, dersom signalet vi sampler inneholder frekvenskomponenter over halve samplingfrekvensen. Disse komponentene speiles nedover i spekteret og danner falske frekvenser.

Bemerk at "samplet signal" også inbefatter signaler som er beregnet ved en matematisk metode internt i en datamaskin. Under beregningene må vi alltid passe på at signalet ikke inneholder komponenter over halve samplingfrekvensen. Disse vil forårsake aliasering. Et enkelt forsøk på en FM-synthesizer (DX 7) vil demonstrere dette. Hent inn en lyd med et rikt spektrum, og spill den øverst på klaviaturet (bruk ev. *Key Transpose* for å komme enda høyere opp). De lyseste komponentene blir speilet nedover, og forvrenger lyden kraftig.

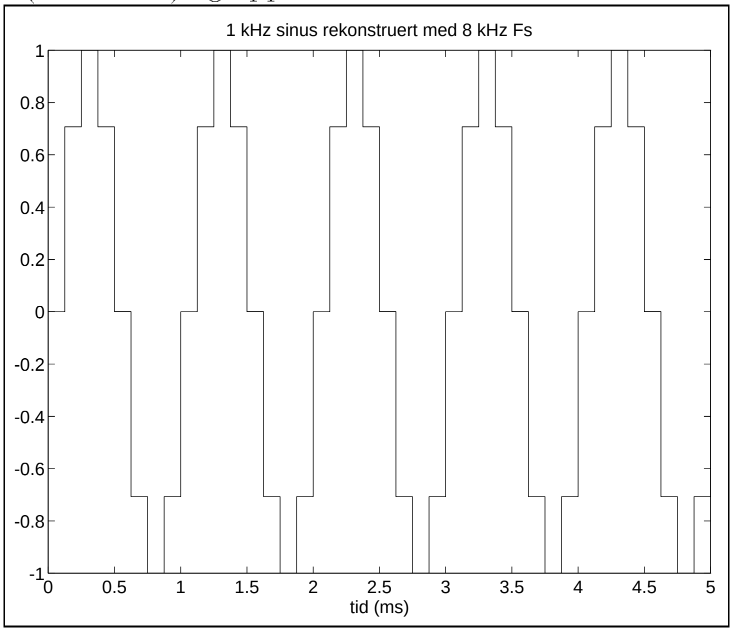
Når vi skal sample en lyd, må vi altså passe på at signalet er så båndbreddebegrenset at aliasering unngås. For å sikre dette, er det vanlig å sette et lavpassfilter før A/D-omformeren, et såkalt *antialiaseringsfilter*. Antialiaseringsfilteret må være så skarpt som mulig, og klippe like under halve samplingfrekvensen.

3 Rekonstruksjon

Nå er altså signalet samplet og lagret digitalt, en tallverdi for hver måling. I dette avsnittet skal vi se på hvordan tallverdiene kan omformes tilbake til et analogt signal. Dette kalles *rekonstruksjon*, og er den omvendte prosess av sampling. En digital/analog-omformer (D/A-omformer) står sentralt i denne prosessen.

En D/A-omformer produserer en spenning på grunnlag av en digital verdi som blir sendt ut fra datamaskinen. For å rekonstruere et analogt signal ut fra digitale verdier, kan vi således sende sekvensen av verdier ut på D/A-omformeren med samme hastighet som signalet ble samplet med. Hvis vi har samplingfrekvens 30 kHz, må vi sende 30 000 tallverdier ut på D/A-omformeren hvert sekund. Vi kan også endre denne hastigheten, med hastighets- og pitch-endring som resultat.

Figuren under viser utgangen fra D/A-omformeren når vi sender ut sekvensen som vi fikk da vi samplet en 1 kHz sinus med 8 kHz samplingfrekvens. Tatt i betraktning at det originale, analoge signalet var en ren sinus, ser jo dette ut som et nokså mislykket resultat. Akustisk vil den "firkantede" formen arte seg som masse bråk og ringing fra Nyquist-frekvensen (her 4 kHz) og oppover.



Ved å betrakte samplingsprosessen som modulasjon med et tog av enhetsimpulser, kan man vise at dette bråket er speilinger av det originale spekteret om Nyquist-frekvensen og heltallsmultipla av denne. For de som måtte ha glede av det, kan dette formaliseres ved at den Fourier-transformerte av signalet x modulert med toget av forsinkede Dirac-impulser er konvolusjonen av den Fourier-transformerte av x med et tilsvarende tog i frekvensdomenet:

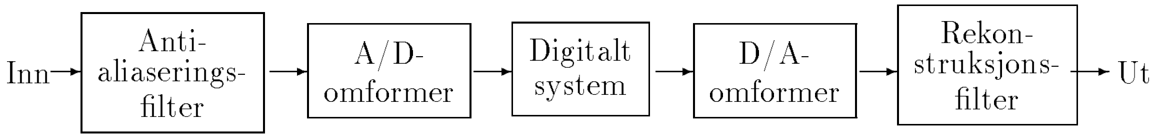
$$\mathcal{F}\left[x(t)\sum_{n=-\infty}^{\infty}\delta(t-nT)\right] = X(f)*\left[\sum_{m=-\infty}^{\infty}\delta(f-mf_s)\right]$$

Det originale spekteret er altså intakt, men har fått masse uønskede komponenter i tillegg, fra Nyquist-frekvensen og oppover. Disse komponentene kan fjernes ved hjelp av et skarpt lavpassfilter, som klipper like under Nyquist-frekvensen. Dette filteret kalles et *rekonstruksjonsfilter*. Etter rekonstruksjonsfilteret skal signalet i prinsippet være perfekt rekonstruert. Trappetrinnene i figuren over vil da være fjernet, og vi vil ha en perfekt sinus igjen. Dette er egentlig et litt merkelig (og for oss svært heldig) resultat. [1]

Det komplette samplings- og rekonstruksjons-systemet er vist på figuren under. Det analoge signalet lavpassfiltreres i antialiaseringsfilteret, slik at vi sikrer oss mot aliasering. Signalet samples og måles så i A/D-omformeren, og lagres og/eller behandles i et digitalt

[1] Dette avsnittet er ikke helt korrekt og fullstendig. En av flere kompliserende faktorer er effekter av pulsbredden fra D/A-omformeren.

system (datamaskin). Tallverdiene sendes videre ut til D/A-omformeren, som produserer en trappetrinn- aktig versjon av det analoge signalet. Trappetrinnene fjernes så i rekonstruksjonsfilteret.



4 Kvantisering

Det finnes enda en viktig faktor som er med på å bestemme lydkvaliteten, nemlig effekten av endelig ordbredde, og dermed avrundingsfeil, i A/D- og D/A-omformerne. Dette gir opphav til *kvantiseringsstøy*.

Ordbredde betyr antall bits (1 eller 0) som brukes for å kode hver sample-verdi. Vi kan til å begynne med tenke oss en 3-bits A/D-omformer. En slik omformer kan skjelne mellom $2^3 = 8$ spenningsnivåer, som kodes med følgende binære verdier (her er fullt utslag 8 volt, og vi regner med bare positive tall):

Spenning inn på omformeren	Kodet verdi	Desimalverdi
Under 1 V	000	0
1-2 V	001	1
2-3 V	010	2
3-4 V	011	3
4-5 V	100	4
5-6 V	101	5
6-7 V	110	6
Over 7 V	111	7

Vi ser her at den amplitydekontinuerte, analoge spenningen, som kan anta enhver verdi mellom 0 og 8 V, blir kvantisert til en av 8 mulige binære verdier. Dette fordi A/D-omformeren har et endelig antall bits (her 3). Kvantisering betyr det samme som at signalet blir gjort amplitydediskret.

Hvilken betydning har så kvantisering for lydkvaliteten? Vi kan finne ut av dette ved å si at kvantisering gir addisjon av støy til det ideelle analoge signalet, som følge av avrundingsfeilene. I eksemplet over ser vi at en virkelig signalverdi på 1.00 V blir kodet som 1. Dermed er signalet målt korrekt. Men en signalverdi på 2.99 V blir kodet som 2, hvilket tilsvarer addisjon av spenningen -0.99 V til det ideelle signalet.

Vi ser at sampling av signalet med vår 3-bits omformer gjør at spenninger fra 0 til -0.99 V blir addert (mikset) til signalet. Hvilken kurveform disse spenningene danner avhenger av det analoge signalets kurveform, men i de fleste tilfeller vil kvantiserings- støyen arte seg som hvit støy. Denne støyen får amplityden 1 V. Sammenlikner vi dette med analogsignalets maksimale utslag, får vi et signal/støy-forhold på

$$\frac{8V}{1V} \approx 18dB.$$

Et signal-støy-forhold på 18 dB holder jo ikke. For å få et bedre resultat, må vi måle spenningen med større nøyaktighet. Dette krever flere bits i omformeren, slik at vi får flere enn 8 nivåer å skjelne mellom. En 8-bits omformer gir f.eks. $2^8 = 256$ forskjellige binærkoder. Kvantiseringsstøyens amplityde blir da $\frac{8V}{256} = \frac{1}{32}V$. Dette gir signal/støy-forhold

$$\frac{8V}{\frac{1}{32}V} \approx 48dB.$$

Mer generelt kan vi sette opp følgende uttrykk for signal/støy- forholdet som følge av kvantiseringen:

$$SN = \frac{A}{\frac{A}{2^n}} = 2^n \approx 6n\ db\ (effekt)$$

der A er maksimalt utslag og n er antall bits. Av dette får vi følgende tommelfingerregel: Signal/støyforholdet som følge av kvantisering, målt i effekt-decibel, er tilnærmet lik 6 ganger antall bits i A/D- og D/A-omformerne. [2]

12-bits omforming gir dermed ca. 72 dB S/N-forhold, mens 16 bits gir ca. 96 dB. Dette er imidlertid rent teoretiske verdier, og unøyaktigheter i omformerne og støy i tilliggende analoge kretser gjør at det reelle S/N-forholdet aldri når opp til denne teoretiske grensen.

Vi kan nå oppsummere hvilken innvirkning sampling har på lydkvaliteten:

- Samplingfrekvensen bestemmer frekvensgangen til systemet. Høyeste målbare frekvens er lavere enn halve samplingfrekvensen (Nyquistfrekvensen). Dersom signalet ikke er tilstrekkelig båndbreddebegrenset, oppstår aliasering.
- Dersom utgangen fra D/A-omformeren ikke passerer et godt rekonstruksjonsfilter, får vi støy på høye frekvenser som følge av spektrale foldinger rundt multipla av Nyquist-frekvensen. [3]
- Signal/støy-forholdet er teoretisk begrenset av ordbredden i omformerne. S/N-forholdet i dB er tilnærmet lik 6 ganger antall bits i omformerne.

[2] Igjen er dette noe forenklet, idet kvantiseringsstøyens energi-innhold gjennomsnittlig er noe mindre enn peak-verdiene vi har sett på. En mer korrekt likning er $SN = (6.02 \cdot n + 1.76)dB$.

[3] Dette kan ofte høres som en ringing på samplere når man transponerer langt ned, fordi rekonstruksjonsfiltrene ikke følger med samplingsfrekvensen nedover. Denne effekten blir ofte feilaktig kalt aliasering.