

Samplerate – mer enn et tall



10. juni 2021

Espen Fjerdingsstad

espefj@imv.uio.no

Masteroppgave 30 poeng

Våren 2021

Innhold

1	Introduksjon	2
1.1	Hva handler denne oppgaven om?	2
1.1.1	Hvorfor samplerate?	3
1.2	Forskningsspørsmål	4
2	Metode	6
2.1	Musikkvitenskapelige bøker og tekster	6
2.2	Musikkvitenskapelige artikler	8
2.3	Fagbøker innen lydteknologi:	9
2.4	Andre bøker og artikler	11
2.5	Avgrensning	12
2.6	Båndbredde og lagring, ikke en uendelig ressurs	13
2.6.1	The Jevons Effect	14
3	Lydteknikk som en fortelling om abstraksjoner	16
3.1	Reisen mot perfekt fidelitet	16
3.1.1	Vil vi noengang ha perfekt gjengivelse?	17
3.1.2	Starten av moderne lydteknikk, papir, voksruller og shellacplater	18
3.1.3	Magnetbånd gjør lydbølgen usynlig igjen	20
3.1.4	Stereo, bevegelsesfriheten blir borte	21
3.1.5	Fra mekanisk til elektronisk forsterkning	23
3.1.6	Radio	23
3.1.7	Digital teknologi, kjeden blir brutt på ny	24
3.1.8	Oppsummering om abstraksjoner	26
4	Samplerate; hva er det?	27
4.1	Hva vil det egentlig si at man <i>tar en prøve</i> ?	27
4.1.1	Droop, jitter, kvantiseringsfeil og dither	29
4.1.2	Foldefeil/Aliasing, intermodulasjon og oversampling	32

4.2	Sampling som en rekke lagringer	35
4.2.1	Bitdybde og dynamisk spenn	38
4.3	Tidlig digitalisering	40
4.3.1	Formater, PCM, DSD og MQA	41
4.3.2	Hvor kommer formatet 44.100 Hz fra?	43
4.3.3	Signalet må innordne seg systemet	44
5	Analyse	46
5.1	Hva er perfekt lyd?	46
5.2	Hva er fordelene med høyere samplerate?	47
5.2.1	Høy samplerate i studio	48
5.2.2	Diminishing returns, lavere avkastning	51
5.3	Hva er fordelene med lavere samplerate?	52
6	Konklusjon	54
6.1	Lyd kan ikke lagres	54
6.2	Avsluttende refleksjoner vedrørende egen oppgave	56
7	Kildeliste	i
8	Tillegg	v

Forord

Det har vært en lang vei å gå for å fullføre denne oppgaven. Flere ganger har jeg vært nær ved å avslutte for tidlig, men ved god hjelp og støtte fra flere, kom jeg i land til slutt. Jeg vil derfor takke min veileder, Hans T. Zeiner Henriksen, for gode råd og innspill. De har vært til god hjelp for å greie å fokusere på denne ene oppgaven.

Videre vil jeg gjerne takke min gode venn og musikk lærer, Rune Nicolaysen. Du har vært en inspirasjon og støtte helt fra jeg startet studiene mine i 2014 og helt til jeg nå har levert denne masteroppgaven. Takk Rune for god stil, som du en gang sa til meg.

Til slutt vil jeg takke medstudenter og kolleger for kollokvier, gjennomlesninger og oppmuntring; Peter, Per Øyvind, Sven Tore, Henrik, Ernesta og Laila.

Jeg ønsker dere alle lykke til med deres studier og prosjekter.

Espen Fjerdingsstad



Figur 1: Reklame fra århundreskifte, (unknown, u.d.); Fair use.

1 Introduksjon

1.1 Hva handler denne oppgaven om?

Denne masteroppgaven er en orienteringsoppgave innen emnet lydteknikk hvor jeg legger til grunnlag andres forskning på området. Jeg stiller meg kritisk til mye av denne tidligere forskningen fordi den gir forskjellige svar på på flere spørsmål, deriblant hvilken samplerate som egner seg best for innspilling av musikk.

Oppgaven skal prøve å belyse et problem jeg har støtt på når jeg har studert musikkteknologi. Det er en myriade av begreper og faguttrykk som aldri blir forklart inngående, men bare illustrert funksjonen av. Videre omhandler det meste av litteraturen rent tekniske aspekter når det jeg har savnet mest er bakgrunnsinformasjon og tanker rundt apparatene og teknologien vi omgir oss med. Enten en gjør lydopptak med en mobiltelefon eller med studiomikrofoner må det gjøres valg, og et av disse valgene er kvaliteten på formatet en velger. Jeg vil i denne oppgaven forsøke å belyse en del aspekter rundt valg av kvalitet, og forsøke å sette mine egne ord på de kriteriene jeg anser som viktige for å fastslå egnet kvalitet satt opp mot ressursbruk.

1.1.1 Hvorfor samplerate?

Gjennom studier av lydteknikk har jeg blitt mer og mer klar over at fremskrittene vi gjør, alltid innebærer en endring av lagringsmediene. Det siste fremskrittet er overgangen fra analog til digital teknologi, og nå skjer lagringen av signalet i form av tall det er mulig å regne på. Denne overgangen starter alltid med å definere en samplerate som det første trinnet for å diskretisere signalet. Samplerate har også en ekvivalent i det kontinuerlige domenet, men der er terminologien frekvensrespons eller frekvensområde. I denne oppgaven vil jeg følge denne tråden som ultima ender opp i det vi i dag kaller for samplerate, men som begynner med en liten nål festet i en membran som tegner en lydbølge på et ark dekket med sot.

Hva samplerate er, kan egentlig oppsummeres iløpet av en setning, det antallet samples¹ et system utfører iløpet av en tidsenhet. Selv om det er en enkel beskrivelse, er begrepet likevel ladet med meninger, både historiske og kvalitetsmessige. Samplerate er et begrep som stammer fra digitaliseringsteknologien, men det har også et historisk innhold. Det er tett relatert til begrepene båndbredde, frekvensrespons og frekvensomfang, som alle er mål på hvilke frekvenser et system kan registrere og lagre. Samplerate er i så måte også et mål på den høyeste frekvensen et digitalt system kan registrere og dermed også et mål på kvaliteten systemet kan produsere.

I det siste har samplerate nærmest blitt et sannhetsbegrep for ekte og autentisk lyd. Ladet med utsagn som «artistens ekte intensjon» eller «slik det låt i studio» er vanlige, og det er god grunn til å bli skeptisk, spesielt når disse argumentene har vært i bruk helt siden lyd først ble kommersialisert og gjort tilgjengelig for massene. Blir lyd alltid bare bedre og bedre eller er det kun idealet om god lyd som forandres? Det som var perfekt lyd igår blir fort utdatert og umoderne, men blir lyden av dårligere kvalitet av den grunn?

Samplerate har helt siden åttitallet vært en del av en terminologi som har ledet mot perfektjon, men vil vi noengang få glede av denne perfektjonen, eller vil dagens teknologi være håpløst akterutseilt om bare noen få år? Diskusjonen om hvorvidt høy-

¹Fra engelsk *sample*; PRØVE

opløselige filer bidrar til bedre kvalitet har vært diskutert lenge uten at det har blitt noen bred enighet om dette. Lenge var det mange som mente at normaloppløsning var så bra som det kunne få blitt ², men med økt kapasitet både på hastigheten og kapasitet av internett, virker det som om trenden har skiftet mot høyopløselige filer for både strømming fra nett og for avspilling fra lokalt lagrede filer.

Allikevel, ser man historisk på det, er det kanskje bare lydutstyret og avspillingsmediene som forandrer seg. Fortellingene om at den perfekte lyd er innen rekkevidde er uforandret. Derfor vil jeg påstå at SAMPLERATE ER MER ENN ET TALL.

1.2 Forskningsspørsmål

Når det er så stor uenighet om hvilken samplerate som egner seg best for musikk, er det på sin plass å stille noen spørsmål rundt hvorfor det er blitt slik. Er det lyd kvalitet som skal veie tyngst når en optimal samplerate skal settes eller er det andre kriterier som må taes med, slike som kostnader ved innkjøp, kostnader ved drift, miljøhensyn for å nevne noen? Hvordan avgjøres lyd kvalitet og hvem skal avgjøre hva som er god kvalitet og hva som er av dårlig? Fraunhofer, utviklerene av MP3, brukte gjerne ingeniørene sine i tidlige lyttetester. Disse kan klassifiseres som eksperter som lytter og de var nyttige i denne fasen, men senere ble det brakt inn andre som sin ekspertise var lytting. De måtte kunne lytte til formatet og ikke innholdet sies det, men selv det bærer med seg innebygde problemer:

Even expert listeners who act like meters [...] represent a broader potential field of users of the technology – something than a universal object. Here we again confront the question of representation: who listens for whom? (Sterne, 2012)[s.166]

Dette som gjaldt for MP3 og det gjelder også for alle andre formater. Når formater som Hi-Res og MQA blir presentert vil de også være nødt til å basere sine produkter på hva de tror forbrukeren vil like, og det finnes ingen standard modell for hvordan

²Fra en test gjort for *Mixonline.com* i 2008 sier Paul Lehrmann «BITS & Hz - Research has shown that for music distribution, 16 Bit @ 44.1 kHz (CD standard) is indistinguishable from 24 Bit @ 192 kHz in a sample of over 550 listeners. In other words, more bits and higher bit-rates are not going to improve the 'quality' of your tracks»(Audio Myths and DAW Wars, 2021)(Lehrmann, 2008)

musikk låter i et studio eller hvilke intensjoner en artist måtte ha hatt da materialet ble spilt inn. På samme måte som Fraunhofer gruppen kom frem til sin modell, har også de andre aktørene kommet frem til sin modell, og da er det naturlig å spørre det samme spørsmålet som Sterne gjør, hvem lytter for hvem?(Sterne, 2012)[s.167]

Alle formater har forskjellige styrker og svakheter, og forskjellige aktører har forskjellige behov, så derfor spør jeg:

Finnes det en optimal samplerate?

2 Metode

Som metode for denne oppgaven har jeg hovedsaklig valgt å drøfte bøker fra musikkvitenskapen, fagbøker om lydteknikk og utvalgte artikler og kapitler fra musikkvitenskapen. Noen andre bøker kommer også inn i blandingen, som jeg har plukket ut fra deres relevans. Her følger en oppsummering og katalogisering av materialet jeg legger til grunn for denne oppgaven:

2.1 Musikkvitenskapelige bøker og tekster

- *Decomposed*, Kyle Devine (2019)
- *Living Stereo*, Paul Théberge et al. (2015)
- *The Audible Past*, Jonathan Sterne (2003)
- *MP3 – The Meaning of a Format*, Jonathan Sterne (2012)
- *Chronopoetics*, Wolfgang Ernst (2016)

Samlingen av litteratur som grunnlag for denne oppgaven begynte med at jeg leste Kyle Devine's *DECOMPOSED* som del av emnet *Political ecology of music* høsten 2020. Siden dengang har henvisninger og sitater i tekstene ledet meg fra den ene boken til den andre. *Decomposed* følger historien om materialbruk i musikkmedier og har således relevans for denne oppgaven i og med at jeg vil skrive om muligheter og begrensninger som ligger iboende i de forskjellige materialene. Sett opp mot samplerate vil jeg se på faktorer som hvor tett informasjon kan lagres på de forskjellige mediene og hvilke frekvensresponser og signal-til-støy ratio de fysisk kan ha.

Etter *Decomposed* har jeg valgt ut Paul Théberge *et.al* *LIVING STEREO*, som er en samling tekster skrevet om fenomenet stereofoni og hvilke kulturelle aspekter overgangen fra mono til stereo førte med seg. Utenom det rent lydmessige, var det også en kjempestor kampanje for rett og slett å få folk til å godta denne nye teknologien. Det var opplæringskampanjer, reklamer rettet mot intellektet og til og med et statlig

prosjekt, da stereolyd i hjemmet ble i USA definert av en del av den amerikanske drømmen³. Jeg ser klare paralleller fra disse kampanjene for stereo da som jeg ser nå for høyoppløselig musikk nå. Bytt ut fenomenet stereofoni med høyoppløselig så ser man at de samme argumentene blir brukt i reklamekampanjene fortsatt. Kan høy samplerate være et fremskritt på linje med utviklingen av stereofonien?

Boken *The Audible Past* av Jonathan Sterne har jeg oppfattet at er en moderne klassiker på områdene innspillingshistorie og kulturen rundt denne historien. Den tar for seg historien fra de tidligste forsøkene på å lagre lyd og til koblinger mellom mellom moderne nedlastninger fra nettet og tidlige forsøk med levende katter som telefoner. Den har i tillegg mange flere eksempler på markedsføringen av tidlige kommersielle formater som voksruller og grammofonplater. I disse markedsføringene finnes verdier som troverdig lyd og autentisitet, akkurat som i dagens markedsføringer. De tidlige formatene påberopte seg de samme kvalitetene som dagens moderne formater gjør, og allikevel hadde de mye lavere frekvensrespons og signal-til-støy forhold. Da mener jeg denne sammenligningen også sier noe om hvor autentisiteten og troverdigheten ligger, og således er viktig å ha med i denne oppgaven.

Den neste boken av Sterne er *MP3 – The Meaning of a Format*. Denne boken er interessant blant annet fordi den tar for seg temaer som koding av signaler, maske-ringeffekter og lytteprøver gjort av ekspertgrupper. Selv om det først og fremst er skrevet for å gjelde MP3 formatet, har tankene til Sterne overføringsverdi til oppgaven jeg skriver om samplerate, fordi samplerate handler også om hvordan vi oppfatter lyd.

Wolfgang Ernst skriver i sin bok *Chronopoetics* om hvordan alle former for signalganger også er en form for lagring eller arkivering i et mediarkeologisk syn. Dette har stor betydning for min oppgave også, da all digitalisering innebærer en samplerate, og denne sampleraten innebærer også en mikrotemporær lagring. Det samme gjelder også for all signalgang, det være seg i ledninger, optiske kabler eller i luft. Jeg

³I 1959 som en del av Nixon og Krustjev's «Kitchen debate» ble medlemmer fra *Institute of High Fidelity Manufacturers* sendt som en delegasjon til *The American Exhibition* i Moskva for å presentere en positiv visjon av USA. (Anderson, 2015)[s.113]

vil gjøre et poeng ut av dette, at når et signal er tatt vekk fra å være akustisk, er det også i lagret tilstand, uavhengig om det er permanent lagret eller temporært lagret.

2.2 Musikkvitenskapelige artikler

- *Analog, digital och sampling – tre centrala begrepp,*
Tellef Kvifte (2012)
- *Vad är egentligen ett fonogram?,* Tore Simonsen (2012)
- *Det klassiske fonogram – et ontologisk dilemma,*
Tore Simonsen (2009)
- *Defining Phonography,* E.W.Rothenbuhler & J.D.Peters (1997)
- *Imperfect sound forever,* Kyle Devine (2013)

Tellef Kvifte har i artikkelen *Analog, digital och sampling* drøftet hvordan terminologien *sampling* har mange forskjellige bruksområder innen moderne musikktenkning. Jeg viser til denne artikkelen når jeg skiller mellom *sampling* gjort i innspillingsøyemed og *sampling* gjort i kreativ, skapende øyemed som for eksempel med en synthesizer. Denne artikkelen er den samme som på engelsk heter *Digital sampling and analogue aesthetics*, men jeg henviser til den svenske oversettelsen fordi det var den er å finne i samme samling som Tore Simonsen artikkel. Samlingen heter *På tal om musikkproduksjon – elva bidrag till ett nytt kunnskapsområde* med Gunnar Ternhag og Johnny Wingstedt i redaksjonen.

Tore Simonsen's *Vad är egentligen ett fonogram?* henviser jeg mye til. Denne artikkelen beskriver hvordan teknologiske fremskritt muliggjør nye manipuleringer av et innspilt materiale. Herfra har jeg hentet mye inspirasjon for å bygge opp min egen teori om at er endringer i lagringsmaterialet som er hovedgrunnen til de nye mulighetene, og også hvordan de samme endringene gir mulighet for høyere samplerate.

Det klassiske fonogram [···] er også skrevet av Simonsen, og er interessant for denne oppgaven, fordi den peker på baksidene ved perfekte digitale kopier, og muligens også

tapet av fonogrammet som kunstart. Hvordan medieprodukter har blitt markedsført som etterligningsapparater fra lydteknologiens opprinnelse. Han viser også til uttalelser fra pianisten Glenn Gould og den amerikanske komponisten Elliott Schwartz' erfaringer med tidlig fonografi, samt siterer medieforsker Philip Auslander.

Defining Phonography: An Experiment in Theory er en artikkel som blir hyppig referert til i de musikkvitenskapelige bøkene nevnt tidligere i kapitlet. Utgitt i 1997, fanger det mye av essensen i overgangen fra fysiske medier til ikke-fysiske. De beskriver også det jeg senere refererer til som brudd med de fysiske aspektene av lydbølgen ved overgangen fra analog til digital lagring (Rothenbuhler & Peters, 1997)[s.252]

Kyle Devine's artikkel *Imperfect sound forever* dreier seg mye om den såkalte LOUDNESS-WAR og også hvordan den har vært med og forme historien om reproduksjon av lyd. Den delen av loudness krigen som kom som følge av digitaliseringen er bare et aspekt av en konkurranse i å bråke mest. En konkurranse som i mine øyne mest sannsynlig startet lenge før sivilisasjonene var begynt. Artikkelen handler også om hvordan pålagte krav, som for eksempel det å spille høyest, alltid vil undergrave lyd kvaliteten og fideliteten. I diskusjonen om samplerate er denne artikkelen aktuell, fordi den peker på det faktum at det konsumentene vil ha, ikke er god lyd, men mye lyd. Det går an å trekke paralleller til digitaliseringen og høy samplerate og si at, det konsumentene vil ha er mange bits, ikke nødvendigvis bra bits.

2.3 Fagbøker innen lydteknologi:

FAGBØKER INNEN LYDPRODUKSJON:

- *The Recording Engineer's Handbook*, Bobby Owsinski (2014)
- *The Mastering Engineer's Handbook*, Bobby Owsinski (2015)
- *The Audio Expert*, Ethan Winer (2018)
- *Mastering Audio*, Bob Katz (2007)
- *Principles of Digital Audio*, Ken C. Pohlmann (2000)

- *Surround Sound*, Tomlinson Holman(2008)
- *Sound for Film and Television*, Tomlinson Holman (2010)
- *Klangen*, Tor Halmrast (2018)

Bobby Owsinskis serie med håndbøker er velrenommerte over hele verden, både for undervisning og som bøker for selvstudier. De er interessante i denne sammenheng fordi de gir en stemme til et vanlig syn når det gjelder sampling, nemlig det at man skal sample så fint som mulig. Dette valget vil jeg gjerne kritisere fordi det gjøres uten en tanke bak, utenom at man gjør det for sikkerhets skyld.

The Audio Expert av Ethan Winer er å anse som et oppslagsverk for alt innen lyd og lydteknikk. Jeg refererer mye til boken i det tekniske kapitlet av denne oppgaven og da spesielt kapittel tre i boken som omhandler hørsel, oppfattelse av lyd og hvordan øret oppfatter artifakter med lyden. Winer er i tillegg en meget kontroversiell person i det profesjonelle lydmiljøet, fordi han er opptatt av å motbevise seiglivede myter om lyd og lydteknikk. Dette har også gjort han upopulær i mange kretser, og han hadde en krangel med lydmikser Eric «Mixerman» Sarafin som gikk fra 2008 til 2018 med offentlige brev og posteringer i forskjellige fora, frem og tilbake i ti år før det roet seg ned. Winer er også kjent for gode lyttetester, både blindtester og vanlige, samt nulltesting for å sjekke differansen mellom to signaler. Han er også motstander av høy samplerate for sampleratens skyld og argumenterer godt for at behandlet romakustikk og gode komponenter i lydkjeden er viktigere enn høy samplerate. Winer er per i dag pensjonert fra profesjonell jobbing i studio, men gjør private prosjekter.

Bob Katz' *Mastering Audio* bruker jeg som oppslagsverk og som en stemme fra en som jobber profesjonelt med lyd på en daglig basis. Han har et meget nyansert syn på samplerate og tar også alltid med kundens behov med i sin omtale om emnet høy/lav samplerate.

Ken Pohlmann har skrevet flere klassikere innen digital lydbehandling, og med sine 700+ sider er *Principles of Digital Audio* en av de mest anerkjente verkene om dette emnet.

Tomlinson Holman's bøker har vært en kilde for illustrasjoner og for en referanse for samplerater brukt i film og TV produksjoner, i tillegg til synspunkter på stereogjengivelse og fantombilde. (Holman, 2008) [s.185]

Klangen er et oppslagsverk skrevet på norsk og samlet av Tor Halmrast. Det har fungert som en referanse på norsk språk

2.4 Andre bøker og artikler

- *Technical Ear Training*, Jason Corey (2010)
- *Zen and the Art of Recording*, Eric Sarafin (2019)
- *Here, There and Everywhere*, Geoff Emerick (2007)

Technical Ear Training omhandler lyttetrening for lydteknikere. Jeg har brukt deler av denne i undersøkelser om hva Jason Corey, som forfatter av faglitteratur, skriver om samplerater.

Eric Sarafin's bok *Zen and The Art of Recording* er med som motstykke til Ethan Winer's *The Audio Expert*. De to har svært forskjellige tilnærminger til faget lydteknikk, og hadde i perioden 2008 til 2018 en debatt gående offentlig i flere forskjellige fora, og dette vil jeg se nærmere på. Sarafin er dessuten også opptatt av samplerates og gir uttalelser i forbindelse med dette. Denne debatten vil jeg omhandle i et senere kapittel.

Lydtekniker Geoff Emerick er mest kjent for sitt arbeid med THE BEATLES, hvor han sammen med bandet og produsent George Martin, la grunnlaget for mye av bandets eksperimenteringer med innspillingsteknikker i studioene til EMI RECORDS i London. Jeg har med denne referansen i historisk øyemed samt for å kaste lys på hva utviklingen av magnetbånd hadde å si for utviklingen av studioteknikk.

2.5 Avgrensning

Det er ikke mulig å skrive om dette emnet uten å utelate mange aspekter. Selv innenfor det begrensede området jeg ser på er det behov for å trekke opp flere grenser. Jeg startet denne oppgaven med et ønske om å se på samplerate som en kulturell konstruksjon og som en utvikling av lagringsmedier og jeg skal nå forsøke å avgrense dette feltet ved å ekskludere visse tråder jeg kunne ha fulgt opp, men som jeg velger å utelate.

Jeg har benyttet meg av en mengde litteratur som kilder fremfor å gjennomføre intervjuer eller forsøk. Det er flere grunner til dette, og jeg vil fremheve kompleksiteten ved og usikkerheten rundt lytteforsøk som noen grunner til å utelate det. Ethan Winer beskriver fremgangsmåter både for *ABX*- og *null*-tester, og begge disse metodene blir regnet som relativt sikre og empiriske, men tester på formater har vært gjort mange ganger før, både av uavhengige forskere og av leverandører av lyd. I tillegg er det en god del kritikk rettet mot disse testen på grunnlag av at de er vanskelige å utføre i naturlige omgivelser og uten stress. Jamfør artikkelen til lydtekniker Allen Farmelo i *TAPEOP MAGAZINE* fra 2012 hvor han kritiserer disse metodene (Farmelo, 2012). Det at de som jobber tett med lyd ofte kritiserer disse testene viser også at de ikke alltid kan gjøres konstruktivt og på en måte som gagnar den som blir testet. Veldig ofte ender det opp i en form for konkurranse hvor den som blir testet ikke må feile for ikke å tape ansikt. Slik argumenterer Eric Sarafin i artikkelen *But How do you Feel?* fra 2017, at for å kunne gi en bedømmelse om et produkt, må en gjøre det om noe en har investert i, enten økonomisk, emosjonelt, prestisje eller alle tre, før resultatet har noen relevans i en profesjonell sammenheng: «[...]if we don't have a stake in the music being successful, we aren't really in a position position to evaluate the sound of it,[...]»(Sarafin, 2017)[s.2].

Fra et forskningsperspektiv er slike tester det eneste en har å gå etter når en skal teste subjektive preferanser, men her er også kompleksiteten stor, slik som den var da Fraunhofergruppen utviklet kodingen for MP3 formatet, men også dem måtte tåle kritikk på grunn av utvalget de hadde av eksperter for å gjennomføre lytte testingen,

som Jonathan Sterne skriver i kapitlet *Of MPEG, measurement and men*: «who listen for whom?» (Sterne, 2012)[s.148]. Rent objektivt sett kan en null-test og blindtester avsløre mye om signalets integritet (Winer, 2018)[s.90]. Jeg anså allikevel at dette falt utenfor rekkevidden av denne oppgaven.

Ett annet aspekt som dukker opp når sampleraten øker og rommene mellom hvert sample minker, men som jeg har utelatt, er *Heisenbergs usikkerhetsprinsipp*. Hvor mye er det mulig å vite om et signal i form av lagret informasjon før det begynner å gå utover integriteten til signalet? Dette var litt for omfattende og avansert til å begi seg i fatt med, men kan være et prosjekt for senere studier. Min egen erfaring på emnet er at det skal veldig lite til av forskyvninger langs tidsaksen før signalet begynner å låte svært kunstig, men for å teste dette kreves lyttetester som beskrevet ovenfor.

Videre har jeg ikke tatt med produktanmeldelser og beskrivelser av spesifikke produkter for å unngå ubalanserte syn med tanke på kommersielle interesser. Jeg har nevnt formatene MQA og HI-Res, men det var i sammenheng med forskjellige måter å modulere et signal på.

Dette er elementer jeg har latt være å ta med i denne oppgaven.

2.6 Båndbredde og lagring, ikke en uendelig ressurs

Selv om miljøhensyn og ressursbruk havner utenfor denne oppgaven, vil jeg allikevel ha med noe om det her i innledningen. Dette vil jeg ha med fordi jeg mener det er viktig for debatten om hva som er den optimale sampleraten.

Ett viktig perspektiv i den sammenheng er hva vi skal gjøre med alle dataene som genereres. Høyere samplerates og større bitdybde stiller ikke bare større krav til levering av datastrømmen, men også til lagringskapasiteten hos leverandørene. Det er lett å tenke på nettskyen som en uendelig lagringsplass, men det er bare i navnet.

Den totale informasjonsmengden som ble skapt, kopiert, lagret og konsumert i verden, passerte 59 ZB (zettabyte=trilliard bytes) i 2020, og vi er inne i en eksponentiell vekst. [...] Mens datamengden eksploderer, flater kapasiteten ut i de eksisterende lagringsløsningene.(Johansen, 2021)

I aprilutgaven av magasinet *Aftenposten INNSIKT* var det en artikkel fra Arctic

World Archive på Svalbard. Her beskriver journalist Kjetil Johansen problemene som kan oppstå når datamengder blir så store at man ikke har plass til å lagre dem noe sted. Selv om nye lagringsmuligheter blir forsket frem, er det dominerende lagringsmediet fortsatt lange magnetbånd. Sett i ett musikkvitenskapelig perspektiv må vi kanskje begynne å bestemme oss for hva som er viktig å arkivere, og hva vi kan la gå.⁴

Selv om fysiske lagringer blir billigere og billigere samt at serverparker kan lagre dataene dine i en nettsky, er det ikke dermed sagt at båndbredde og lagringskapasitet er en uendelig ressurs. I tillegg til materialer som går med til produksjonen av lagringsmedier, er lagringsmedier både lokalt og sentralt kraftkrevende å ha i drift. Begreper som *strømming* og *nettsky* er gode på å lede tankene mot noe godt og naturlig, og de skjuler den teknologiske virkeligheten som ligger bak i form av nedlegging av milevis med kabler og fiberoptikk og

2.6.1 The Jevons Effect

Det sies at høyoppløselige filformater bringer lytteren nærmere artistenes ekte intensjoner, men dette er det eldste salgsargumentet for å selge musikk som finnes. Ser en litt videre i problemet, kan dette behovet for høyoppløselig musikk like gjerne være en respons på den stadige utbyggingen av nettkapasitet, som det er en søken etter den optimale lyd kvaliteten.

Kyle Devine skriver i sin bok *Decomposed* om hvordan Jevon's effekt, oppkalt etter økonomen William Stanley Jevons, legger til grunn at effektivisering av en teknologi ikke fører til innsparinger, men snarere til et økt forbruk. Jevons bruker drivstoff forbruk i sin analyse, og sier «It is wholly a confusion of ideas to suppose that the economical use of fuel is equivalent to a diminished consumption. The very contrary is the truth».(Devine, 2019)[s.208]

Dette paradokset vil gjelde like mye for ressursene nettkapasitet og lagring av data, som det gjorde for bruken av drivstoff.

⁴Når det gjelder musikk er mesteparten av informasjonen lagret fra rundt 50 Hz opp til rundt 10 kHz, et piano har frekvensomfang fra ca 28 hz til 3850 Hz.

The Jevons effect

It would be reasonable to expect that squeezing more energy out of a given resource should make for less overall consumption of that resource. Yet the opposite is usually true. This is the Jevons effect: greater efficiency in resource usage is actually followed by a greater consumption of that resource. [ibid.s.33]

Sett fra et miljøhensyn er det da slett ikke sikkert at høyoppløselige signaler er gunstig.

*There is no abstract art. You must
always start with something.
Afterwards you can remove all
traces of reality.*

Pablo Picasso

3 Lydteknikk som en fortelling om abstraksjoner

3.1 Reisen mot perfekt fidelitet

Historien om lydteknikk er ofte blitt presentert som et kontinuerlig narrativ om frem-skrutt om hvordan lyden progressivt har blitt bedre og mer realistisk for hver nye metode og hvert nye format som har blitt tilgjengelig. Denne tolkningen er selvfølgelig gunstig for produsenter av musikk og selgere av musikkutstyr, men kan det være en objektiv sannhet i det også?

Dette kapittelet vil også inneholde min egen tolkning av historien om lydteknikk. Hvordan hvert skritt fremover også vil innebære å legge fra seg elementer fra den gamle teknologien. Jeg vil ved hjelp av utvalgte eksempler vise hvordan en rute mot perfekt fidelitet og naturtro gjengivelse samtidig bringer oss vekk fra lydens naturlige opphav. Dette er et fenomen jeg mener er universelt innen all lydteknologi og kanskje dermed for all annen forskning også. Men, som Jonathan Sterne påpeker, er historien om fidelitet, en historie som gjerne blir fortalt med slutten først:

Histories of sound fidelity usually begins at the end, with the achievement of perfect fidelity and flawless sound reproduction. Narratives of technological change and the transformation of technical specifications are folded back into an aesthetic and technological telos: the latest technological innovation equals the 'best-sounding' or 'perfect' sound reproduction. The progress narrative is ultimately untenable: the transformation of practises and technologies stands in for a narrative of vanishing mediation, where sources and copies move ever closer together until they are identical. (Sterne, 2003) [s.222]

Hva vil skje hvis perfekt gjengivelse kan oppnås? Vil vi da bare kunne ha perfekte kopier og ingen originaler lenger? Det digitale domenet gir muligheter for å lage

kopier som er identiske i lagringen av bits, men er dette noen garanti for at de kan gjenskape perfekt fidelitet?

3.1.1 Vil vi noengang ha perfekt gjengivelse?

Det kan vel kalles et paradoks at vi på en og samme tid både nærmer oss og fjerner oss fra det vi ønsker å kalle *naturlig*. På en annen måte er dette problemet bare en utskifting av verdier, et sett med verdier hvor lyd kvalitet bare er en verdi blant mange. Det vil være naivt å tro at alle fremskritt blir drevet frem av krav om lyd kvalitet. Det som er drivende faktorer er like ofte eller oftere faktorer som, mangel på råmaterialer, muligheter for masseproduksjon, bekvemmelighetshensyn for forbrukeren, og andre liknende faktorer. Et format som voksrullen ble dominert av grammofonplaten, ikke så mye av lyd kvalitet, som på grunn av tilgjengelighet. Grammofonplaten kunne enkelt masseproduseres ved hjelp av stemping, mens en voksrull måtte til å begynne med spilles inn i sanntid. Senere ble det utviklet metoder for å kjemisk fremstille voksruller, men selv det var ikke nok til å konkurrere med grammofonplatene på pris og tilgjengelighet.

Det var også et annet element med i denne kampen om formatene. Noe som vi er berørt av den dag i dag, og som nok mest sannsynlig stammer tilbake til tiden lenge før grammofonspillere og moderne lydteknikk, nemlig forsøket på å spille høyere enn konkurrentene. Grammofonspilleren var i stand til å spille ved et mye høyere volum enn fonografen, og dette ga dem også en stor fordel. Voksrullene var simpelthen ikke egnet til å lage så kraftige spor som måtte til for å sette så store svingninger i bevegelse. På den tiden var alle lydrepoduserende apparater helmekaniske og alle lyder måtte forsterkes opp akustisk ved hjelp av fokusering og projisering. Naturlig nok gikk også et økt volum ut over fideliteten, og grammofonspillerene fikk også den karakteristiske sterke gjengivelsen i mellomtoneregisteret som vi forbinder med dette formatet ⁵. LOUDNESS-WAR er således ikke et moderne fenomen, men hadde sitt

⁵Grammofonspillere kunne spille så høyt, at det ble vanlig å putte noe i lydtrakten for å dempe volumet. Det var mulig å bytte trakt og/eller lydhode for å oppnå det samme, men det var enklere å dytte noe ned i trakten. Dette er også muligens opphavet til det engelske uttrykket «Put a sock in it.»

opphav for lenge siden.

3.1.2 Starten av moderne lydteknikk, papir, voksruller og shellacplater

I denne oppgaven skal jeg begynne med begynnelsen og jobbe fremover. Som en merkelig kuriositet, er det ikke lenger Edison som laget det første lydopptaket lenger. Organisasjonen *First Sounds*⁶, har lyktes i å trekke ut lyd fra franskmannen Édouard-Léon Scott de Martinville sine lydskrift eksperimenter med *fonoautografen* (Firstsounds, u.d.). Fonoautografen var i utgangspunktet kun ment for å visualisere lyd og tale, ikke for avspilling, men med moderne optisk lesing og datarekonstruksjon lykkedes det å bringe frem verdens første lydopptak, en fransk barnesang. Med denne inngripenden i historiens gang ble startskuddet for lydteknologi flyttet fra 1877 til 1860.

Den tidlige teknologien fungerte ved å benytte en trakt som endte i en membran. I denne membranen var det festet en registreringsnål som overførte vibrasjonene i luften til svingninger på et medium. Denne teknologien var basert på forskning på den menneskelige hørsel, og de første eksperimenterne til Thomas Edison og hans assistent Watson var også basert på de indre delene av et ekte menneskeøre, den såkalte *Ear-Phonograph*. Disse tidlige eksperimenterne var heller ikke beregnet bare på musikk, men også som hjelpemidler. De var tenkt som et verktøy for opplæring i tale for døve og tunghørte. Slik at de kunne se skriften av talen sin og korrigere seg selv etter skriften av talen til en lærer. I tillegg ble fonografen markedsført som en diktafon og transkriberingshjelpemiddel-hjelpemiddel.

For å kunne reproducere lyd måtte det en ny abstraksjon til. Det var gjort tidligere forsøk med å reproducere stemmer ved å etterligne et strupehode (*automata*), men det var først da Edison begynte å bruke en membran (*tympanisk*) til høyttaler at utviklingen begynte å gå fremover. Nå var behovet for et strupeliknende apparat borte og øret hadde således blitt delegert begge oppgavene, både innspilling og avspilling.

men dette er ikke bekreftet.

⁶<http://www.firstsounds.org/sounds>

(Jonathan) Faber copied the movement of the vocal organs, Edison studied a vibrating diaphragm, and produced the action of the eardrum when acted upon by the vibration *caused* by the vocal organ. (Sterne, 2003) [s.71]

Der Faber søkte en årsak, søkte Edison en effekt, og selv om automata-konstruksjonen var mye mer avansert enn membranens forholdsvis enkle tympaniske funksjon, var Edison's tilnærming mye mer egnet til å gjenskape alle lyder, ikke bare stemmer. Ved å mikrofons teknologi til både innspilling og avspilling, hadde Edison nå en anvendbar metode for å konstruere *fonografen*.

Tett beslektet til fonografen finnes *grammofonen* og teknikkene for hvordan både grammofonplater og voksruller fungerte var nærmest identiske. Begge formatene var helt mekaniske både i drift og forsterkning. Ved innspilling er det en diger trakt fokuserer lyden ned til et innspillingshode. Vibrasjonene fra lyden blir fanget opp i en membran som sender lyden videre til en inngraveringsnål som igjen lagrer et bilde eller en representasjon av lydbølgen på rullen/platen.

Lägg märke till att en akustisk inspelning på många sätt är den mest korrekta återgivningen av en akustisk händelse som tänkas kan, på grunn av sina begränsningar. Balansen mellan instrumenten på fonogrammet var densamme som inspelningstratten hörde, ljudförloppet från början till slut var en sann avbildning av det som försiggick under inspelningen. (Simonsen, 2012) [s.110]

Denne tanken som Simonsen beskriver er den samme tanken jeg skal forfølge og også utdype underveis. Et poeng til med voksruller og shellacplater som jeg skal utdype, er at man kan se avtegningen av lydbølgen. Det er ikke mulig å gjøre etterarbeid som nivåjustering, redigering, tonebalanse og lignende, etter at lyden er spilt inn⁷. Dette kommer først med senere teknologiske fremskritt. Det at det fortsatt er mulig å kjøpe disse formatene via bruktmarkedet, gjør voksrullen til verdens eldste format som fortsatt er i bruk. I og med at fonografen var konstruert for opptak, er det fullt mulig å benytte disse apparatene til å gjøre nye opptak den dag i dag.

Frekvensomfanget på disse formatene var svært begrenset. Deg har ikke lyktes meg å finne noen estimeringer på dette, men jeg vil gjette på rundt 200Hz til 3000Hz. Det som går over dette høres ut som å være støy i mediet blandet med litt musikalsk

⁷Med fonoautografen var det i prinsippet mulig å redigere ved å klippe i papirstrimmelen og lime den sammen igjen hvis det var ønskelig.

innhold, men dette blir litt gjetting fra min side. For signal-til-støy virker dette til å være en del dårligere enn en dårlig kassett, så da vil jeg gjette på rundt 30dB, eller en 3-5 bits, avhengig av materiale og kvalitet på innspillingen. Allikevel er dette vanskelig å fastslå uten ekstra målinger og analyser, men det kan uansett anslås at det ikke låter veldig realistisk sett etter moderne standarder. Disse tidlige mediene har også en spesiell følsomhetskurve, og dette vanskeliggjør sammenlikningen med dagens lydutstyr ytterligere, så mine antakelser her er, i beste fall bare antakelser.

3.1.3 Magnetbånd gjør lydbølgen usynlig igjen

Magnetbånd åpnet for en helt ny æra innen innspillingsteknikken og muligheter i et musikkstudio. Der hvor en før var avhengig av ett eneste opptak, hvor alt måtte klaffe ned til minste detalj, kunne en med magnetbånd få flere sjanser og etterhvert kunne man ta opptak av ett og ett instrument for senere å sette det hele sammen til et enhetlig musikkstykke. Temporaliteten og en kausal rekkefølge var temmet, men det kom med en pris.

Med muligheten til å etterarbeide et opptak, var det følgelig også mange som gjorde det. Innen popmusikk kunne en gjøre det ganske åpenlyst, men for andre sjangere som for eksempel klassisk, var det tanken om det komplette og sammenhengende opptaket som dominerte;

[...] Den här delen av klassisk musikproduktion har aldrig fått någon stor uppmärksamhet. Myten om att långa, helst kompletta tagningar ger bättre flyt och sammanhang än inspelningar med många klipp är svår att bekämpa. Den kanadensiska pianisten Glenn Gould var på 1970-talet ganska ensam om att hävda att «good splices build good lines». (Simonsen, 2012) [s.112]

Det var ikke bare fordeler med å kunne gå tilbake å fikse på ting, dekke over eller gjøre opptak på nytt. Innspillingsprosessene hadde nå potensiale til å mangedobles i kompleksitet og grensene for hva som var autentisk og hva som var konstruert ble stadig visket bort. I sin biografi gir Beatles tekniker Geoff Emerick et øyeblikksbilde fra sin karrière:

By 1967, eight-track tape recorders were already starting to appear in professional studios [...] yet the entire *Sgt. Pepper* album was recorded in four-track. In

short order, sixteen-track and then twenty-four would become standard recording tools, [...] George Martin has said in many interviews that *Pepper* wouldn't have been as good had it been recorded in twenty-four-track, and I completely agree. It was because of those very limitations that we were put on the spot [...], and that was part of the magic of the album. You had to put the right echo on, the right EQ, the right signal processing; the playing had to be right, the vocal had to be right. (Emerick & Massey, 2002)[s.191]

Det neste jeg vil påpeke er at ett element til av lydopptaket var abstrahert bort. Nå var det det synlige elementet av lydbølgen som var borte. Til gjengjeld hadde man nå fått tilgang til å manipulere tiden. Med et enkelt kuttebord, barberblad og litt tape, kunne opptak som var gjort i forskjellige tidsrom spleises sammen, man kunne spille det raskere og langsommere, lage sløyfer av lyder som gjentok seg i det uendelige, spille perfekte kopier baklengs, og dette var bare noen av mulighetene som åpnet seg, nå når tidsdomenet kunne beherskes.

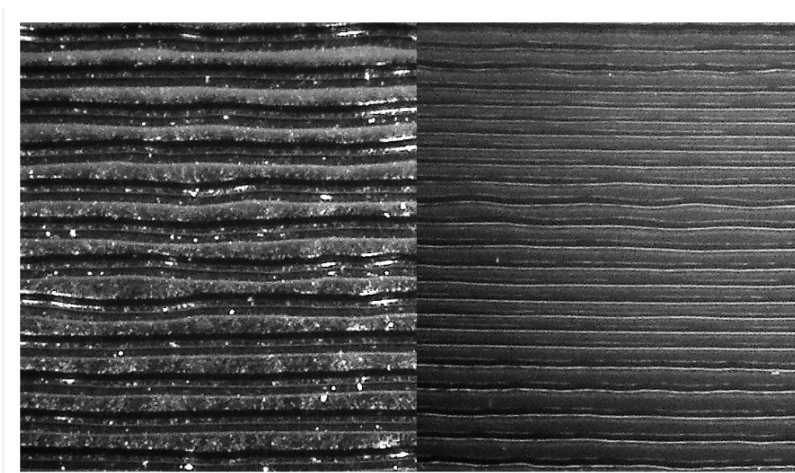
3.1.4 Stereo, bevegelsesfriheten blir borte

Det kan virke som litt ute av sin plass å bringe opp stereofoni i utviklingen av formater, men overgangen fra mono til stereo tar faktisk bort et element også. Det er ikke mulig å oppnå noen stereoeffekt, uten å faktisk befinne seg på en veldig spesiell plass mellom høyttalerene, det såkalte «sweet spot». Utenfor dette punktet vil ikke stereobildet manifesteres. Stereobildet er også kalt «phantom image» og er en effekt av at lyden kommer strømmet ut fra et sted midt mellom høyttalerene.

Det var krefter både for og imot denne nye teknologien, og blant annet en frustrert leser av magasinet *High Fidelity & Audiocraft* uttalte sin frustrasjon i leerspalten slik:

Well, I've had it. I welcomed LPs and the expensive equipment for playing them as the genuine advances they were. During the last decade I've junked a fortune in 78s and slowly built up a library of some four hundred LPs, with no regrets [...], but I love music and I get it (in both ears!) [...] so now comes something with the name of «stereo» that is supposed to bring the concert hall to us at home. (Moore, 1959)

Stereoeffekten var enda ikke etablert som noen norm, men allikevel presset musikkindustrien på for at det skulle bli det. Det konsumerende publikum måtte læres opp til å bruke denne nye teknologien og det ble gjort ved hjelp av en rekke over-



Figur 2: Til venstre ser vi rillene på en grammofonplate, til høyre en vinylplate. Simonsen (2021)

talende informasjonskampanjer. Det bærer likheter til hvordan dagens konsumenter læres opp i høyoppløselige formater. Fra å være en sosial aktivitet i mono hvor alle kunne få mer eller mindre den samme lyden, var nå stereo blitt en seriøs og solitær utøvelse av lytting.

Denne forestillingen om den ensomme mann som kjenner sitt lydanlegg holder seg den dag i dag bare med ny terminologi. Paradokset er at de aller fleste lytterom ikke fungerer så godt til stereolytting, mye på grunn av dårlig akustisk behandling av rommet og suboptimal plassering av høyttalere og lytteposisjon. Det kan hende monoinnspillinger kunne gjort et comeback. Geoff Emerick skriver i sin biografi at i 1967 var det fortsatt monomiksene som fikk den beste behandlingen i mikserrommet (Emerick & Massey, 2002) [s.170].

Vi fikk et stereobilde i bytte mot friheten til å bevege oss rundt mens vi lytter. Dette kan virke som en ubetydelig handel i fremskrittet mot en naturlig lyd, men overgangen var alt annet enn naturlig.

Yet stereo, which has been handed down to us as a natural and inevitable stage in the inexorable march of «high fidelity,» was anything but a smooth step in that familiar narrative of progress. Indeed, the massive shift in just a few short years from mono to stereo was fraught with uncertainty and even resistance, marked not only by uncritical celebrations but also by serious confusion and vigorous debate over its apparent virtues. (Grajeda, 2015) [s.39]

3.1.5 Fra mekanisk til elektronisk forsterkning

Den mekaniske overføringen av lydbølger fra et lagringsmedium til lydbølger i luft på grammofon og voksrull var helt mekanisk og uten annen form for forsterkning enn rent akustisk. Dette endret seg med elektronisk forsterkning og senere når LP tok over som lagringsmedium, var elektronisk forsterkning den eneste muligheten for å få lyden ut i rommet. Man kunne fortsatt høre lyden uten elektrisitet hvis en la øret helt inntil stiftene, men dette signaler var ikke sterkt nok til at det egnet seg for akustisk forsterkning.

Det som overfører signalet som ligger på en LP plate er en meget fin stift i ene enden som agiterer et magnetfelt i den andre enden, og bevegelsen til stiftene tilsvarer bevegelsen til apparatet som inngraverte signalet i rillene på platen. Denne overgangen gjorde at stiftene ikke lenger trengte så stort trykk og sammen med at materialet i en LP er mere holdbart, kunne musikk lagret på LP også være mere holdbart og tåle flere gjennomspillinger.

Som vises i figur 2 ser vi at både hos LP og grammofonplater er lydølgen godt gjenkjennelig, men en LP har mye finere spor, og dermed ikke samme utgangspunkt for å sette i gang svingninger som blir til lydbølger. Dette sporet er hva som menes når Rothenbuhler og Peters refererer til når de beskriver en analog kjede som ubrutt, mens en digital 3 er brutt.

3.1.6 Radio

Midt oppe i all denne lokale lagringen er det lett å glemme kanskje det viktigste av alle medier, nemlig radioen. Her er sampleraten styrt av veldig mange faktorer, men allikevel var gode FM-sendinger tidligere å betrakte som et av de beste høyfidelitets-signalene en kunne spille gjennom et lokalt lydanlegg.

En godt konstruert FM-mottaker kan gjengi stereoprogrammer av beste kvalitet ved at signalene kommer direkte fra mikrofoner istedet for fra opptaks/avspillingsutstyr. (Hjelm & Haave, 1978)[s.53]

Med radioen ble musikk gjort tilgjengelig for mange flere. Stemmer og musikk

kunne forflytte seg fritt mellom tid og sted, fra konsertsaler og inn i stuene. Det er nesten så en kan høre ekkoet fra Adorno når Rothenbuhler og Peters påpeker «*Radio turns the transcendent achievement of Beethoven into kitsch.*»(Rothenbuhler & Peters, 1997)[s.257]

Det innebar også en abstraksjon og diskretisering av signalet da sendinger gikk over til å bli sendt som FM (frekvens modulasjon) fra å tidligere å ha vært sendt som AM (amplitude modulasjon). I tillegg til å ha en hjelpebølge som bærer signalet innebærer frekvensmoduleringen et brudd med tiden og sendes som små pakker med informasjon istedenfor en kontinuerlig strøm. Denne teknologien ble ikke sett på som diskretisering på samme måte som vi nå ser på en strøm av data, men mediearkeolog Wolfgang Ernst fremviser slektsskapet her:

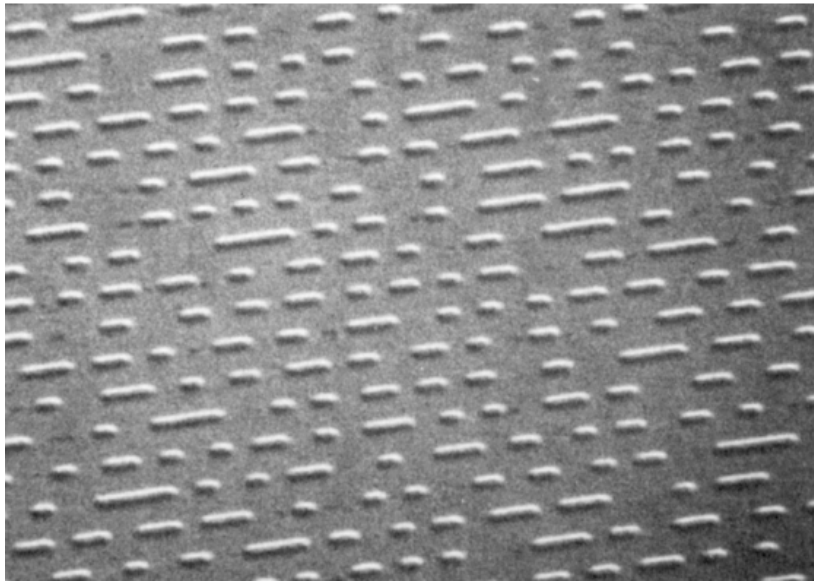
With classical amplitude modulation (AM), radio signals were completely continuous; frequency modulation (FM), on the other hand, introduced the low-frequency signal (speech or music) as a *time event* into the formerly equidistant carrier frequency–chronotechnical immediacy. However, an entire media-epistemic world lies between continuously modulated signals and a signal form that consists of periodically consecutive pulses.(Ernst, 2016)[s.183]

Det viser seg at denne overgangen mellom kontinuerlig og diskret er glidende og det samme gjelder for de fleste digitale systemer. Det er ved lagring at det er et markant skille mellom de to domene, men straks et signal skal begynne å bevege seg, kan det ikke være fullstendig diskret lenger, men må også inneha kontinuerlige egenskaper.

3.1.7 Digital teknologi, kjeden blir brutt på ny

Overflaten på en CD, som er avbildet i figur 3, viser mønsteret av et digitalt signal. Dette signaliserer et brudd med den analoge inngraveringen av et lydspor som finnes på LP og grammofonplater, men dette betyr ikke at ikke CD teknologien har analoge elementer ved seg.

Av alle brudd med signalgangen er nok overgangen fra analog til digital lagring det nest største, etter det opprinnelige bruddet med tid og sted som ble begått i det øyeblikk lyd ble registrert og lagret. Det opprinnelige bruddet hvor en stemme kunne



Figur 3: Bilde av informasjonen på en CD.(Fair use)

separeres fra avsenderen var naturlig nok det mest dramatiske, *Its uncanniness consists in its ability to suppress absence, to span the chasm between source and addressee that death and distance once seemed to make impossible.* (Sandberg, 1991).

Analog recording has a physical relation to the sound recorded and digital does not. [...] Analog recording is, in essence, a trace of the phenomenon such that analog playback, in principle, can be a reproduction. Digital recording is a measurement of the phenomenon and so digital playback must be a reconstruction. This simple distinction is rich with implications. (Rothenbuhler & Peters, 1997) [s.252]

Hva Rothenbuhler og Peters her fremhever, som en del av en lengre meditasjon over fonogrammer, er en distinkt forskjell mellom analog og digital lagring. På platen er det fortsatt et vinglete spor som er en tegning av lydbølgen, mens en digital fil kun inneholder data som i prinsippet kan være hva som helst. Den binære informasjonene av en datafil bærer ingen distinkte ytre forskjeller fra en hvilken som helst annen datafil, og dermed kunne dataene like gjerne vært et bilde eller en tekst.

Jeg vil argumentere for at det er ikke her kjeden blir brutt for første gang, men at kjeden her har blitt brutt *nok en gang*. Første bruddet var helt tilbake til separasjonen av stemmen fra taleren, og dette bruddet var mye mer signifikant enn dette siste bruddet. Nå er vi forberedt og det skal mye mere til for å lure noen i dag med et

opptak av en stemme enn det var for 150 år siden. Selv om selve lyd kvaliteten på mediene på overflaten har blitt bedre jo flere brudd fra kilden det har, sliter dagens musikkteknologi med et autensitetsproblem.

3.1.8 Oppsummering om abstraksjoner

Jeg ønsket å skrive dette kapitlet om abstraksjoner fordi jeg mener det har sammenheng med hvordan vi kan se på samplereater også. Med de tidlige formatene var alltid frekvensresponsen begrenset av fysiske lover, i motsetning til digitale signaler som bare består av tallverdier. Ved å fjerne alle rester av signalet og bare ta vare på en digital representasjon av det, kan nye metoder tas i bruk for å etterbehandle det.

Når signalet først er lagret i digital form er det helt frikoplet fra fysiske hemninger. Det er ingen egenvekt som kan hindre opptakssystemet i å ta opp alle frekvensene som blir sendt til det. Det er ikke dermed sagt at noen mikrofon greier å registrere alle frekvensene som opptaksutstyret kan bruke. For eksempel har en mikrofon fysiske begrensninger i form av membran og elektroniske komponenter.

Ved å se på abstraksjoner som årsaken til at utviklingen går fremover, kan kanskje en vei videre fra den digitale teknologien også vise seg.

4 Samplerate; hva er det?

We do not describe the world we see, we see the world we can describe.

Descartes

Sampling; ETYMOLOGI engelsk, til *sample*, 'prøve, utvalg'

Rate; ETYMOLOGI av latin *rata pars*, 'fastsatt del'

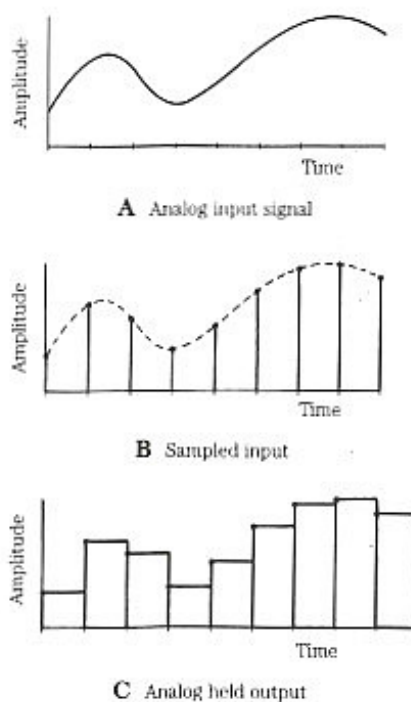
Begrepet SAMPLING har sin etymologiske opprinnelse fra latin og engelsk, ([Rate](#), 2018), ([Sampling](#), 2018)

4.1 Hva vil det egentlig si at man tar en prøve?

Innen samplingsteori vil det å ta en prøve innebære at man henter ut en verdi som beskriver energinivået i en lydbølge på et visst tidspunkt, og lagrer denne verdien som et diskret tall. Dette gjøres ved et system som kan omsette den målte verdien til en binær verdi, og et slikt system kalles gjerne gjerne en ANALOG-TIL-DIGITAL KONVERTER, og er å finne som mellomledd i alle lydkort som bindeledd mellom det analoge og det digitale domenet.

Selve prosessen som skjer er meget innviklet og avhenger av en rekke tidsavhengige mindre ledd. En måte å ta prøver på er ved såkalt SAMPLE-AND-HOLD metode. Som navnet tilsier, er dette en metode hvor et signal blir lest av og holdt i et mellomlager, inntil det kan lagres som en digital verdi. Hvor lenge signalet blir holdt, avhenger av samplingfrekvensen til systemet, og systemet må holde den mellomlagrede verdien til datamaskinen skifter register, noe som skjer i takt med samplingfrekvensen. I et lite øyeblikk, rett etter at verdien er digitalt lagret, og rett før en ny verdi lagres i mellomlageret, befinner det samme signalet seg både i det analoge og det digitale domenet, og samplingen er så nøyaktig som den kan få blitt. Etter noen flere sykluser er det analoge signalet borte, mens dets digitale representasjon ligger lagret i dataminnnet, og her kan det ligge lagret upåvirket av tiden.

Conceptually, an S/H⁸ circuit is a capacitor and a switch. The circuit tracks the analog signal until the sample command causes the digital switch to isolate the capacitor from the signal; the capacitor holds this analog voltage during A/D⁹ conversion. (Pohlmann, 2000)[s.56]



Figur 4: En grafisk fremstilling av SAMPLE-AND-HOLD prosessen. (Pohlmann, 2000)[s.56]

I figur 4 vises en grafisk fremstilling av sample-and-hold prosessen. Øverst er det analoge signalet som en kontinuerlig bølge. Deretter blir det tatt samples punktvis langs denne bølgen og i dette stadiet har ikke punktet noen utstrekning i tid. Det nederste bildet illustrerer hvordan den samlede verdien holdes i en kondensator i påvente av at signalet leses digitalt. Lesingen kan blant annet skje ved hjelp av et nettverk av identiske elektriske motstander. Legg også merke til hvordan denne fremstillingen minner om fremstillinger av hvordan et digitalt signal ser ut, mens en trappetrinnsmodell i virkeligheten viser hvordan et digitalt signal er lagret.

The concept of chronology, which is so familiar in our culture, depends on the

⁸Pohlmann har forkortet SAMPLE-AND-HOLD til S/H

⁹Analog-to-Digital

technical processing of binary signals. The object of switching algebra is logical and thus not time-dependent connection, but the relations between input and output signals in delay and memory circuits as well as frequency multiplication are especially relevant in their temporality. (Ernst, 2016) [s.67]

Her påpeker Ernst den dualistiske naturen til en datamaskin når den blir brukt i en mediesammenheng. Selve signalet er temporært, men alle logiske prosesser i datamaskinen er ikke tidsavhengige. I en ideell situasjon, hvor alle delene av systemet har de ønskede karakteristikker, og klokkefrekvensen kan holdes helt nøyaktig, vil denne formen for sampling kunne være en likefrem og enkel operasjon. I virkeligheten vil alle elektriske komponenter ha små avvik fra den oppgitte verdien og klokkefrekvensen kan inneha bittesmå avvik, noe som igjen kan gi grobunn for støy på signalet.

4.1.1 Droop, jitter, kvantiseringsfeil og dither

Digital sampling kommer med en hel rekke problemer og utfordringer som er unikt for et digitalt format. I tittelen på dette avsnittet har jeg ramset opp noen få, men digitalisering er så mye mer avansert enn noe annet format, så mulighetene for komplikasjoner er mange flere. Richard Feynman beskriver i sin selvbiografi hvordan han lærte å reparere radioer bare ved å åpne dem og kikke inni, Kittler søkte å forstå radioer ved å observere dem fra utsiden, digital signalbehandling må studeres både innenfra og utenfra.

DROOP: I en analog-til-digital konverterer er det to elementer. Ett element som tar vare på en verdi i form av en spenning, og en annen som leser av denne spenningen som en binær verdi. Når et signal når kondensatorene (les. elementet) som skal «holde» på spenningen til et system av motstander (les. det andre elementet) går det en liten tid fra signalet leses til det kan konverteres. I dette lille tidsrommet kan spenningen i kondensatoren synke litt og man får en litt lavere verdi enn det skulle være (Pohlmann, 2000) [s.58]. Dette kalles *droop* og er et fenomen som kan gjøres mindre med bedre komponenter, høyere samplerate og korrigeringsalgoritmer.

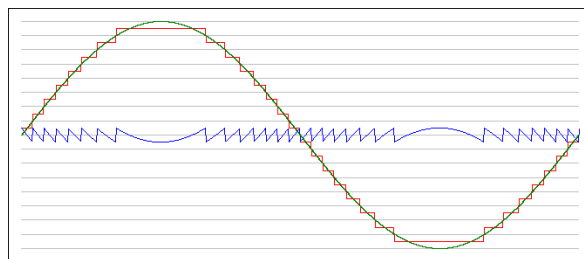
JITTER: Dette er en feil som vil oppstå på tidsaksen når klokkefrekvensen ikke er nøyaktig nok. Unøyaktigheten fører til at små forskyvinger på bølgeformen som vil kunne manifesteres seg som støy. I det analoge domenet kalles tidsfeilene for *wow* og *flutter*. Disse gir små endringer i tonehøyde, men jitter fluktuerer så raskt opp og ned at feilen oppfattes som støy. Selv om dette er et reellt problem, er det likevel stor uenighet om hvor alvorlig man bør behandle dette. Ethan Winer har gjort utallige tester på lyd og fidelitet og han skriver:

Digital recorders and soundcards have a type of timing error called *jitter*, but the pitch deviations are so rapid that they instead manifest themselves as added noise. With all modern digital audio, jitter is so soft compared to the music that it's almost inaudible. (Winer, 2018)[s.47]

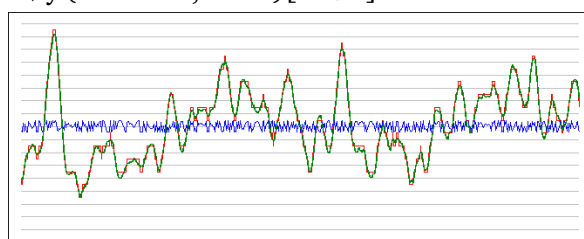
KVANTISERINGSFEIL: I en samplingsprosess må signalet kvantiseres før det kan lagres. Dette gjøres ved at det kontinuerlige signalet gis den nærmeste verdien det har det har passert. Dette fører til at biter av signalet kappes av og dette vil igjen produsere det som kalles for kvantiseringsfeil/støy. Kvantiseringsstøy for rene sinus-toner er veldig påtrengende, men for et signal med vanlig musikk vil støyen være tilnæringsvis lik vanlig hvit støy.

DITHER: Dither er lavnivåstøy som tilføres et digitalisert signal, som regel når det skal avkortes til en lavere bitdybde. Det brukes også for å bekjempe andre typer støy, slik som kvantiseringsstøy i samplingsprosessen. Det sies at ditherstøy bør være randomisert for å ha best mulig effekt, men en hvilket som helst støykilde kan gjøre jobben. Problemet er at ekte random er vanskelig å oppdrive. Det viktigste er at ditherstøyen ikke er relatert til signalet. Problemet er neglisjerbart for de fleste komplekse lydkilder, slik som musikk, men kan være et problem ved svake lydkilder ¹⁰. Ditherstøy kan hjelpe et signal til å låte renere, men som Ethan Winer påpeker er det

¹⁰With large amplitude complex signals, there is little correlation between the signal and the quantization error; thus the error is random and perceptually similar to white noise. With low-level signals, the character of the error changes as it becomes correlated to the signal, and potentially audible distortion results. (Pohlmann, 2000)



(a) Kvantiseringsfeil av en ren sinusbølge vil gi en lett gjenkjennbar og mekanisk støy. (Halmrast, 2018) [s.375]



(b) For et normalt signal, vil kvantiseringsstøyen være spredt utover et mye større frekvensspekter. [ibid.]

Figur 5: Kvantiseringsfeil/støy.

aldri hovedgrunnen til at noen mikser ikke låter så bra:

[...] Some engineers even argue over which type of dither is best, claiming that this algorithm is more airy or more full sounding than that one, and so forth. But just because everyone believes this, does that make it true? To be clear, using dither is never a bad thing, and it can reduce distortion on soft material recorded at very low levels. [...] Not using dither is never the reason an amateur's mixes sound bad. (Winer, 2018) [s.77]

For å understreke poenget sitt, har Winer publisert en test av åtte lydfiler av en sang som heter *Lullaby*, hvorav noen er påført dither mens andre er direkte avkortet [ibid.]. Lytteren kan utføre sin egen blindtest for selv å avgjøre effekten av dither.

We played several different short sections of Grekim's tune this way, and neither of us heard any difference between truncated and dithered. [...] Even with the volume cranked on a very soft passage we still heard no difference. Even I was a little surprised by that! (Winer, u.d.-a)

Ovennevnte eksempel viser at det er mye som tyder på at viktigheten av dither er sterkt overdrevet, men det er blitt standard prosedyre og så godt som alle lydteknikere tilfører dither ved nedsampling. Dette reiser spørsmål om hvorfor det gjøres, men

heldigvis, dithersignalet er uansett så svakt at det kan vanskelig gjøre noen skade på signalet uansett. Det passer fint inn i utsagnet «Do no harm!», som ofte blir brukt om miksing og mastringsprosessen. Som mastringsingeniør Ian Shephard så fint sa det i sin podkast *The Mastering Show, episode 21*, «Stop dithering about dither and just... dither!»



Figur 6: Go ahead, share this as much as you like (I.Shephard)

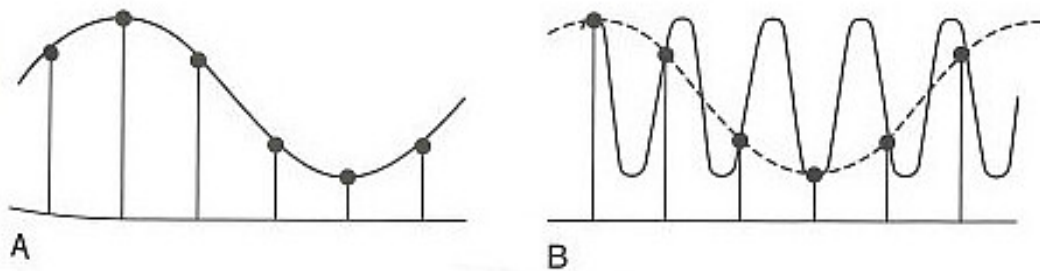
For å være på den sikre siden....

4.1.2 Foldefeil/Aliasing, intermodulasjon og oversampling

FOLDEFEIL/ALIASING: Aliasing ble først brukt om et fenomen ved radiomottaking med superheterodyne radiomottakere. Superheterodynmottaker er en fagbetegnelse på en helt vanlig radiomottaker. I en slik mottaker kunne man få et ekstra *signal* eller *alias*, som interfererte med bærebølgen i signalet en ønsket å lytte på, og således ga en gjenkjennbar effekt (Aliasing, 2021). Også innen billedbehandling kan aliasing oppstå. I stillfotografi kan det gi en effekt som også er kjent som *moiré*, som er et gjenkjennbart interferensmønster som kan oppstå ved billedbehandling. Et annet eksempel er kjerrehjul på gamle cowboyfilmer som ser ut til å bevege seg baklengs. Disse eksemplene viser at aliasing ikke er et fenomen bare for digital sampling, men at det er et fenomen som kan oppstå overalt hvor oppløsningen på mottakermediet er mindre enn oppløsningen på signalet som skal registreres og lagres.

Derfor må et signal som skal samples, også ha begrenset båndbredde, slik at ikke

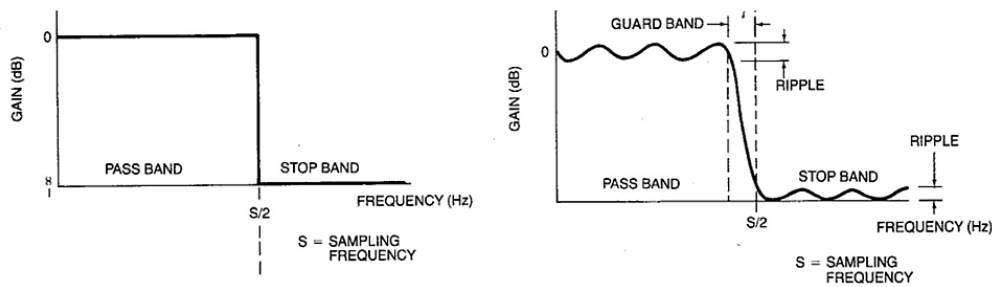
analog-til-digital konvertereren mottar høyere frekvenser enn den kan håndtere. For digital sampling er dette beskrevet i *samplingsteoremet*, som fastslår at for å kunne beskrive et signal må samplingsfrekvensen være dobbelt så høy som den høyeste frekvensen en ønsker å sample.



Figur 7: Figuren illustrerer hva som skjer når det ikke er nok samples til å få med to samples per svingning. (Fair use); (Holman, 2010) [s.51]

På figur 7 ser vi tydelig hvordan et nytt signal oppstår når sampleraten er for lav til å håndtere det innkommende signalet. Aliasing kan forhindres ved hjelp av filtre som filtrerer ut alt frekvensinnhold over den kritiske båndbredden, og det er sjelden vi hører aliasing i musikkproduksjon takket være dette. Fordi aliasing ikke er harmonisk relatert til signalet, kan det være en meget skjæmmende artifakt hvis det ikke blir kontrollert. Rundt den kritiske båndbredden vil ikke aliasingen være veldig prominent, men jo høyere over den kritiske båndbredden signalet er, jo lengre ned i det hørbare spekteret vil også aliasingen kunne forekomme. Aliasing kan enkelt demonstreres ganske tydelig ved opptak av sinussveip, og da låter aliaset som en svak fuglekvitring/plystrelyd.

En måte å effektivt forhindre aliasing på, er ved å filtrere ut alt det uønskede frekvensinnholdet før signalet blir digitalisert. Dette gjøres ved hjelp av lavpass filtrering



Figur 8: Til venstre er øyeblikkelig stopp illustrert. Dette er kun oppnåelig i teorien, så til venstre er en illustrasjon av et lavpass filter i praksis. (Fair use);(Pohlmann, 2000)[s.53]

(fig.8). Filtrene kan være enten analoge, digitale eller en kombinasjon av de to, men felles for dem alle er at gode filtre er kostbart å fremstille, og jo skarpere de skal være, jo større krav setter det også til konstruksjonen av dem. I moderne sampling forsøker man å flytte problemet lenger vekk fra det hørbare spekteret ved at man sampler med en frekvens som tilsvarer en, to eller noen ganger tre oktaver over normal oppløsning.

INTERMODULERT FORVRENGNING: En egenskap hos instrumenter som kan forveksles med aliasing. Intermodulert forvrengning har mange av de samme egenskapene som aliasing, men det vil oppstå av forskjellige grunner. Derfor kan de to ofte forveksles, men IMF er ikke et fenomen som kommer på grunn av digitaliseringen. Intermodulert forvrengning kommer som regel i par med harmonisk forvrengning, men der hvor harmonisk forvrengning kan være ønskelig eller intermodulasjon sjelden det. IMF inntreffer når to eller flere frekvenser spilles samtidig og instrumentet er egnet til å skape rene toner. Da kan summen og forskjellen mellom danne nye svingninger som ikke er harmonisk relatert til grunnfrekvensene. Eksempelet under er en A ved 440 Hz og C# ved 277 Hz spilt sammen. (Winer, 2018)[s.46]

$$Sum : 440Hz + 277Hz = 717Hz$$

$$Difference : 440Hz - 277Hz = 163Hz$$

Her ser vi at frekvensen av summen blir 717 Hz som er halvveis mellom en F og F# og 163 Hz som er litt under en E. Ingen av disse tonene er relatert til A, heller ikke til noen standard stemming, så en intermodulasjonsforvrengning som dette vil kunne produsere ubehagelig frekvensinnhold, og visse forsterkere reagerer sterkere på dette. En løsning for å rette på dette er å øke innholdet av den harmoniske forvrengningen ved å bruke effekter som SATURATION, OVERDRIVE eller lignende. [ibid.]

OVERSAMPLING: For å gjøre filtreringen ved den kritiske frekvensen lettere å kontrollere, er det mulig å oversample signalet. Alle filtre har sideeffekter, og selv kvalitetsfiltre kan ikke kutte frekvenser brått, uten at artifakter som ripler i passbåndet, støy eller faseforskyvinger kan oppstå, men jo bedre kvalitet filteret har jo mindre artifakter har det også, men prisen blir også deretter. For å bøte på dette kan man benytte seg av oversampling for å gi filteret flere frekvenser over det hørbare spektrum å jobbe med. Dermed kan Q-kvaliteten¹¹ gjøres lavere for å gjøre kurven på filteret slakere og dermed får filteret flere frekvenser over det hørbare spektrum å jobbe med. Ved å flytte stoppbåndet lenger opp i frekvensområdet unngår man mange problemer. Oversampling gjøres i det digitale domenet ved å benytte seg av DELTA-SIGMA MODULERING¹², og først etter filtreringen, desimere signalet til PCM¹³, et format som datamaskinen kan lese¹⁴. En annen måte å gjøre dette på, er ved å sample i høyere oppløsning i PCM format, og denne metoden har blitt mer populær i det siste, med tilgjengeligheten av raskere datamaskiner og større harddisker.

4.2 Sampling som en rekke lagringer

Istedenfor å tenke på sampling som en kjede med hendelser kan man tenke på det som en rekke med lagringer. For at datamaskinen skal kalles fullstendig diskret, må

¹¹Q(quality) er båndbredden til et filter og angir forholdet mellom midtfrekvensen frekvensbåndbredden på karakteristikken 3dB under midtfrekvensen (Halmrast, 2018)[s.431]

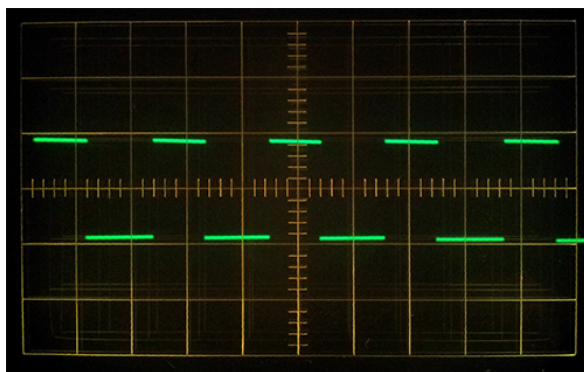
¹²Delta-Sigma er et høyfrekvent 1 bits representasjon av signalet.

¹³Pulse Code Modulation.

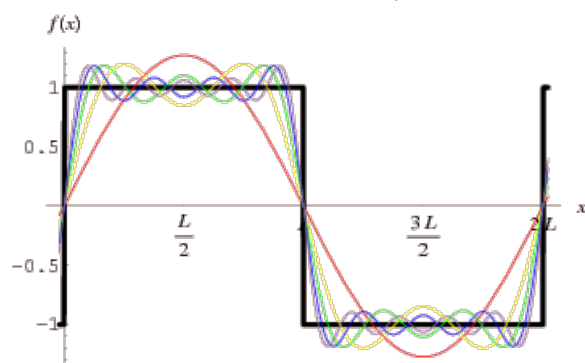
¹⁴Det finnes noen få DAW som kan jobbe med DSD signaler, men de er sjeldne og dyre.

også enhver tilstand kunne være fullstendig repeterbar. I virkeligheten skjer ikke alle hendelser samtidig i en datamaskin, og i de små tidsforskyvingene, vil det være momenter av usikkerhet hvor systemet må ta valg som kan være forskjellige fra gang til gang.

Selv etter at signalet er digitalisert, vil det fortsatt være rester av dets analoge opphav i signalet. En puls i en datamaskin blir ofte fremstilte som en firkantbølge, men selv den består av en rekke sinusbølger. Eneste gangene signalet ikke befinner seg i bølgeform, er når det står i ro, det vil si er når det er i en lagret tilstand.



(a) På et oscilloskop tegnes ikke rampene inn. (Fair use; AntiMath) (Mihay, 2021)

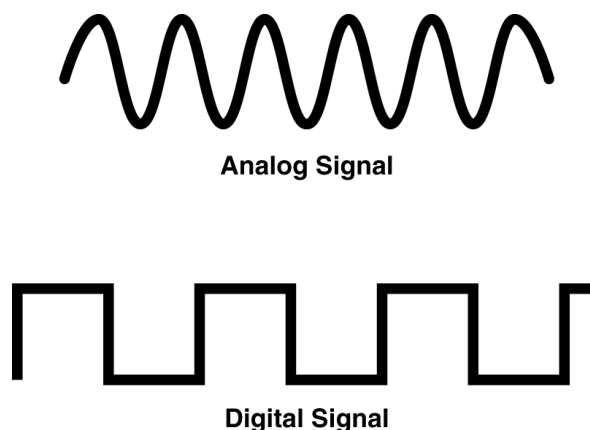


(b) I en illustrasjon kan vi tegne inn flere iterasjoner. (Fair use; Wolfram Alpha) (Weisstein, 2021)

Figur 9: Forskjellen mellom representasjon og illustrasjon.

Som vi ser i figur 9 skjuler det seg mer bak en firkantbølge enn det som vises på denne innstillingen på et oscilloskop. Dette går igjen i for illustrasjoner av firkantbølger og pulser, hvor byggestenene som er sinusbølger, ikke blir tatt med av praktiske

hensyn. Dette er underforstått for en oppvakt lesere som kjenner til bølgeteori, men for en som ikke kjenner til teorien vil det se ut som en bølge kan gå i rette vinkler.



Figur 10: Illustrasjon av et analogt og et digitalt signal (Fair use); ([What are the types of signals, 2021](#))

På figur 10 ser vi to bølgeformer. Den øverste illustrerer en sinusbølge, og den kan representeres med to verdier, amplitude og frekvens. Den nederste er en firkantbølge, men som vist i figur 9 består jo en slik bølge av en rekke sinusbølger, nærmere bestemt en grunntone og en rekke av oddetalls overtoner. Jo skarpere kanter en vil ha, jo flere overtoner og høyere frekvenser kreves for å tegne den, og for å tegne helt rette vinkler kreves en uendelig mengde av sinusbølger i en overtonerekke. Jeg ønsker å påpeke dette da slike figurer finnes på filtre, equalizere og illustrasjoner av samplede signaler, hvor firkantbølger blir tegnet som trappetrinn.

Wolfgang Ernst skriver i boken *Chronopoetica* om hvordan alle signaler i prinsippet også er i en lagret tilstand helt fra de går over fra å være bølger i luft. Dermed er det ikke bare når signalet ligger lagret på et permanent lagringsmedium at det kan regnes som å være lagret, men heller helt fra signalet blir omformet til et format vi kan bearbeide. Her vil jeg skille mellom en digital lagring og en analog lagring. Jeg liker å betrakte prosesseringen som skjer i en datamaskin som en serie med adskilte hendelser, og hver av disse hendelsene kan igjen ses på som separate tilstander. Hver av disse tilstandene er mulig å rekonstruere helt perfekt fordi de er diskrete. Et kontinuerlig signal går også gjennom forskjellige tilstander, men her kan ikke de

forskjellige hendelsene rekonstrueres perfekt, fordi overgangene er glidende mellom hendelsene. Allikevel vil en datamaskin ha elementer av kontinuerlig signalgang, fordi ikke alt i en datamaskin ikke skjer samtidig. Selv med hastigheten til elektroner, tar det en viss tid for signalene å gå gjennom alle prosessene internt i datamaskinen. Alt kan ikke skje samtidig, og derfor er det ikke lenger sikkert at hver tilstand kan rekonstrueres lenger, og systemet er ikke lenger fullstendig diskret.

Technological media always take place in the temporal dimension, regardless of whether they are understood through epistemological reflection. Technical constellations are only operative when actualized in time. Technical constellations are only operative when actualized in time. Time-critical processes – such as delicate electronic synchronization between image senders and receivers («television») and the exact orchestration of binary construct instruction cycles («computer») – occur in ELECTRO-TECHNICAL (commonly called «analog») and TECHNO-MATHEMATICAL (commonly called «digital») media.
(egne uthevinger) (Ernst, 2016)[s.3]

Den delikate synkroniseringen Ernst beskriver belyser også et viktig poeng i min diskusjon. Istedet for å tenke på et samplet signal som låst til en samplerate eller bitdybde, må dette differensieres i flere grader. I stedet for å tenke på en samplerate som 44.1 kHz, 48 kHz, 88.2 kHz, 96 kHz og så videre kan man heller se på hvilket frekvensomfang man får ut av systemet. Hvis den delikate balansen i systemet forrykkes, vil dette gå utover frekvensomfanget og signal-til-støy ratioen i form av introdusert støy.

4.2.1 Bitdybde og dynamisk spenn

Det er vanlig å spesifisere bitdybde som heltall innen spesifikasjoner for datautstyr, og i nyere tid har det også blitt vanlig å gjøre dette som multipler av 8. For et lydsignal har en bitdybde typisk representert et dynamisk spenn eller et støygulv, men hvis vi heller ser på det dynamiske spennet som en representant for en bitdybde, kan vi beregne den ekvivalente bitdybden til et ikke-digitalt apparat også.

I figur 11 vises den beregnede SIGNAL-TIL-STØY ratioen for en gitt bitdybde. Det er vanlig å tenke på bitdybde som noe som kun angår inngangssignalet, men jeg mener det er vel så viktig å fokusere på hva en får ut av systemet. Det er beregnet

Number of bits	Dynamic range
8	48
10	60
12	72
14	84
16	96
18	108
20	120
22	132
24	144

Figur 11: Bitdybde og signal-til-støy relasjon; (Wadell, 2013)[s.65]

at hver ekstra bit gir omtrent 6 dB lavere støygulv, og en slik beregning er det som ligger til grunn for å angi det dynamiske spennet for en bitdybde. Det vil da være mer hensiktsmessig å relatere det dynamiske spennet til en bitdybde istedet for å si at en bitdybde gir et vist dynamisk spenn. På den måten tas hele systemet i betraktning og ikke bare den teoretisk mulige signal-til-støy ratioen.

For en typisk kassettpiller er det dynamiske spennet på rundt 40 til 50 dB. For en AKAI CS-35 er det 46 dB, og leser vi i tabellen 11 ser vi at det tilsvarer en bitdybde på litt under 7 bits. En litt dårligere kassettpiller kan ha helt ned i 35dB dynamisk spenn, og da vil den tilsvarende bitdybden tilsvare litt under 6 bits. Dette er spesifikasjonene fra fabrikken, og med litt bruk og slitasje, ville nok avstanden til støygulvet blitt betraktelig mindre. Chris «Monty» Montgomery utdyper dette i en informasjonsvideo, og der viser han også til platespillere og spolebåndspillere, med henholdsvis omtrent 11dB for platespiller og 14dB for spolebåndspiller i dynamisk spenn (Montgomery, u.d.).

Ved å følge denne metoden kan en få en større forståelse av hva digitale signaler egentlig er, og at forskjellen mellom analogt og digitalt egentlig bare ligger i hvordan signalet blir lagret.

Hvis man kompenserer for potensielle støykilder, vil også et digitalt system trenge korrigerings. Dette viser at det er umulig å regne effektive bits ut av et system basert på antall bits systemet opererer med i lagret tilstand.

Table 5-1 Number of Bits versus Dynamic Range

<i>Effective number of bits</i>	<i>Dynamic range, dB*</i>
16	93
17	99
18	105
19	111
20	117
21	123
22	129
23	135
24	141

*Includes the effect of triangular probability density amplitude function dither, preventing quantization distortion and noise modulation; this dither adds 3 dB to the noise floor to prevent such problems.

Figur 12: Bitdybde kontra dynamisk spenn når normale støykilder blir kompensert for; (Holman, 2008)[s.142]

4.3 Tidlig digitalisering

Det er en vanlig misoppfattelse at digitalisering innebærer en elektronisk datamaskin, men dette er ingen nødvendighet. Digitale systemer kan forekomme i mange forskjellige varianter. Et godt eksempel på dette er det såkalte *PLAYER PIANO*, et automatisk piano ofte kjent fra western filmer og lignende. Dette pianoet hadde musikken lagret på en papirrull med huller i, hvor hver hull tilsvarte en tonehøyde. Det var et marked for preprogrammerte ruller, og disse var det som regel en ekte pianist som hadde programmert ved å spille dem inn.

Komponist Conlon Nancarrow brukte dette formatet til sine komposisjoner for *player piano*. Istedenfor å spille inn komposisjonene med tangentene, anvendte han istedet kalkulator og målebånd for å regne ut avstanden mellom hullene, og laget hullene utenfor pianoet. På den måten skapte han musikk i et digitalt format lenge før datamaskinen var realisert for musikkbruk.

Det er mange systemer som kan kalles digitale, ikke bare datasystemer som vi tenker på i dag. Et eldgammelt system med to verdier er for eksempel tenning av varder, hvor et bål betyr noe og ikke et bål betyr noe annet. Videre er for eksempel morsesignaler, alfabeter og notesystemer alle eksempler på digitale kommunikasjonsmetoder. Språk og utøving av musikk er igjen eksempler på analoge kommunikasjonsmetoder.

Selv om prinsippet om digitalisering kan bli enkelt forklart, er digital teknologi alt annet enn enkel. Det er en svært sofistikert teknologi som tøyser grensene for hva som er mulig å få til på mange felter, både innen utvikling og produksjon. Som Ken C.Pohlmann skriver i innledningen til første kapittel i *Principles of Digital Audio*:

Although the underlying concepts have been well understood since the 1920s, commercialization of digital audio did not begin until the 1970s, simply because theory had to wait 50 years for technology to catch up. The complexity of digital audio is all the more reason to begin the discussion with the basics.(Pohlmann, 2000)[s.1]

Nå, over hundre år etter konseptualiseringen av digital audio, står fortsatt det meste uløst. Det er ingen enhetlig teori for hvordan lyd best kan konverteres og det er heller ingen konsensus om hvilken samplerate som best egner seg for lyd. Kompleksiteten til digital teknologi er fortsatt overveldende, og noe av det som er mest frustrerende, er hvordan en teknologi som i utgangspunktet er så enkel, kan være så uforståelig.

4.3.1 Formater, PCM, DSD og MQA

I teorien er det nærmest et endeløst nummer av forskjellige teknikker som kunne vært brukt for å kode et signal. Ikke alle egner seg like godt, *amplitude modulasjon(AM)* og *frekvens modulasjon(FM)* er eksempler på en type signalmodulasjon som blir brukt for radiosendinger. I tillegg til MP3 er det tre formater som kan regnes som normale for musikk i dag, PCM, DSD og MQA.

Modulation is nothing more than a means of encoding information for the purpose of transmission or storage.(Pohlmann, 2000)[s.49]

PULSKODEMODULASJON (PCM) PCM er den vanligste form for signalmodulasjon, og det er også den mest brukte kodingen for musikk. Det er også den eldste av av de tre og ble utviklet i 1937 for bruk innen telefoniteknikk(Pohlmann, 2000)[s.51]. Ved å representere et analogt signal med en pulskode, kan et binært tall brukes til å beskrive

den pulsens amplitude. Senere ved rekonstruksjonen kan prosessen reverseres, og den analoge bølgeformen rekonstrueres ut fra pulskoden.

Pulskodemodulasjon er altså et gammelt format, men det har stadig blitt utviklet, og brukes nå i så og si alt av DAW¹⁵, på CD, DVD, DAT med mere. PCM har den fordel at det tegner et komplett øyeblikksbilde av lydbølgen, og dermed egner det seg meget godt til redigering og andre former for manipulering, hvor man vil fjerne tiden fra signalgangen. Bakdelen er at det er plasskrevende og at det oppdateres langsommere enn andre formater. PCM er det eldste og mest utbredte av alle formater, med unntak av MP3¹⁶.

DIRECT STREAM DIGITAL (DSD): For å bøte på begrensninger med CD-formatet, utviklet SONY og Phillips etterhvert formatet Super Audio CD(SACD), som benytter seg av DSD(Direct Stream Digital) modulasjon.

When it was launched in 1982, the CD was rightly heralded as a data carrier of immensely capacity. However, over time the CD's 650 Mbyte capacity seemed increasingly small. In particular, the CD is insufficient for the large file sizes and high bit rates required by surround sound and high sampling rate audio. (Pohlmann, 2000)[s.295]

For å forlenge levetiden for CD-formatet, ble SACD lansert som en oppfølger. Formatet benyttet seg av DSD modulering, og dette er en annen type koding enn PCM modulering. Ved DSD er bitdybden bare 1 bit, men til gjengjeld en sampleraten helt oppe i MHz området. DSD benytter seg av en form for SIGMA-DELTA modulasjon, som ekstremt hurtig måler summen av endringene i lydbølgen og lagrer dette som en 1-bits datastrøm. Denne typen modulering er avhengig av at datastrømmen er i bevegelse hele tiden, og derfor egner dette formatet seg dårlig for bruk i redigeringsprogramvare. Til redigering egner PCM og dets «øyeblikksbilder» seg bedre. Det finnes riktignok redigeringsprogrammer for DSD også, men det er for spesielle applikasjoner.

Grunnet den høye samplerate som var mulig å oppnå med DSD, ble digitalutgan-

¹⁵Digital Audio Workstation

¹⁶MP3 er egentlig ikke et format, men en komprimering i form av persektiv koding av et lydsignal.

gen på SACD-spillere blokkert for alt annet enn dedikerte digital-til-analog konvertere, slik at datastrømmen ble beskyttet mot piratkopiering. Dette var et resultat av at plateselskapene ikke skulle miste kontroll over masteropptakene sine, og grunnet de dyre lisensene SONY og Phillips måtte betale til plateselskapene, ble også SACD et meget dyrt format å produsere. Til tross for dette er det stadig nye utgivelser på formatet og brukte SACD har også høy andrehandsverdi.

MASTER QUALITY AUTHENTICATED (MQA): Et tredje format som har kommet på markedet er MQA, og det er i vinden som aldri før. Det benytter seg hverken av DSD eller PCM modulasjon, men er et proprietært format som strekker seg fra koder til dekker, og i likhet med DSD, vil heller ikke MQA la bitstrømmen avleses. I tillegg kan autentiseringen gjøres av eller i samarbeid med artisten, og etter at filen er autentisert og kodet, kan den heller ikke gjøres om uten å bryte autentiseringen. Dette gjør at strømmetjenester og andre kanaler ikke kan endre materialet for å passe det inn i sin katalog, ei heller legge til eq, komprimering eller andre ting, noe som er mulig med PCM formatet.

En stor ulempe med MQA formatet er dets avhengighet av dedikerte digital-til-analog konvertere for å spille den høyoppløselige delen av filen. Filen kommer pakket slik at tre oppløsninger er tilgjengelig i den samme filen, og dette er meget kontroversielt, spesielt med tanke på at konseptet er såpass forskjellig fra PCM. Dette har ført til at formatet har mange motstandere.

4.3.2 Hvor kommer formatet 44.100 Hz fra?

Som samplerate var lenge 44.100 Hz det mest vanlige formatet, men hvor kommer dette tallet fra? Som en verdi er tallet 44100 er flere ting, blant annet:

$$2^2 \cdot 3^2 \cdot 5^2 \cdot 7^2 = 44.100$$

Produktet av kvadratet av de fire første primtallene utgjør tallet 44.100. Dette gir

mange heltalls faktoriseringer, og det kan ha gjort utregninger lettere å utføre for en datamaskin, i og med at en datamaskin håndterer binære verdier, men det er en litt løs teori. Ken Pohlmann har beskrevet formatet mer utfyllende:

In a monolithic digital world, there would be only one sampling frequency. In practise, the world is populated by many different sampling frequencies. Although 44.1 kHz and 48 kHz are the most common, 32 kHz is used in many broadcast applications [...] Many wonder why 44.1 kHz was selected as a standard sampling frequency. Originally the only tape recorder capable of handling the high bandwidth requirements of digital audio signals were video recorders. (Pohlmann, 2000)[s.458]

Det viser seg altså at samplertaten 44.100 Hz gjorde at man kunne bruke videoapparater bergegnert for både den europeiske standarden PAL og den amerikanske NTSC for å lagre digitale filer samplet i denne frekvensen. Videobånd var også det eneste formatet som kunne holde noen større mengde data.

4.3.3 Signalet må innordne seg systemet

Et lite poeng som kom med digital sampling, var at nå var det signalet som måtte finne seg en plass og innordne seg i lagringsmediet, i motsetning til analog teknologi, hvor lagringsmediet ble formet etter signalet. Slik som en platekutter avtegner lydbølgen i vinylen, er det også lagringsmediet som formes etter signalet som skal lagres der. Med digital sampling er lagringen allerede delt opp i faste plasser, og signalet må tilordnes den verdien i lagringsmediet som ligger tettest opp til verdien signalet har. I praksis har ikke dette noen betydning, for oppløsningen i en datamaskin er allikevel mye høyere enn det som trengs for å lagre lyden perfekt, men fra et mediearkeologisk synspunkt er det absolutt av interesse.

I en veldig kjent forelesning gjort i 1985 av fysikeren og nobelprisvinneren Richard Feynman, beskriver han datamaskinen som et veldig effektivt arkiv. Dette arkivet blir drevet av en utrolig rask, men fullstendig dum arkivar. Forelesningen er kalt *Idiosyncratic Thinking; Computer Heuristic Lecture* og er et godt eksempel på den utrolig enkle oppbyggingen av en datamaskin, men allikevel hvor komplekst et datasystem kan være (Feynman, 1985). Hvis man ønsker å følge denne tråden videre, kan man

se på matematiker John Horton Conway's *Game of Life*, hvor et spill med tre enkle regler, kan brukes til å konstruere datamaskiner som selv spiller game of life¹⁷

En av de største fordelene med digitalisering, er at vi nå kan redigere lyden uten å modifisere originalopptaket. Ved hjelp av metadata kan vi ved beskrivelser av operasjonene som vi ønsker å gjøre utføre endringer uten å endre selve lydfilen. Det vil være tilstrekkelig å gjøre disse endringene ved bruk av metadata. Det kan være klipping, redigering eller pålegg av effekter, eller alt annet som er mulig å med opptaket.

¹⁷Game of Life er Turing komplett, og samtidig så enkelt at det kan programmeres i et lite program. (Det var et spill som pleide å være inkludert i instruksjonsbøker for hjemmedatamaskiner tidligere, da gjerne skrevet i BASIC)

5 Analyse

Blue light rain, unbroken chain

Looking for familiar faces

In an empty window pane

Unbroken Chain; Grateful Dead

(Lesh, Peterson)

5.1 Hva er perfekt lyd?

Utgangspunktet for denne oppgaven var å finne ut om det kan finnes en optimal samplerate for musikk, den endte opp med å bli en sammenholdning mellom forskjellige lagringsformater. Hvordan skal en kunne finne en optimal samplerate hvis det ikke er mulig å finne ut hva som er god lyd.

Jeg har fulgt med i debatter angående høyoppløselig musikk og digitalisering av musikk generelt, og det jeg merket var hvordan fokus alltid rettes mot antall bits inn og kun i mindre grad på hvor mange bits man får ut.

Lyd er bevegelse og lyd kan aldri stå i ro og det er kun i nedskrevet form at en beskrivelse av lyden kan lagres. Det spiller ingen rolle om lagringen består av et innriss på en plate eller om det er noter til et musikkstykke. Bevegelsen er stanset og det kreves et arbeid for å sette lydbølgene i gang igjen. Likeledes kreves det et samtykke mellom lytteren og teknikken om at lyden en hører er autentisk og riktig. Kriteriene for et slikt samtykke endres med tiden, og dermed er det sagt at det som låt ekte og autentisk for hundre år siden, kan låte kunstig og fremmed idag. Som Jonathan Sterne påpeker: «After 1878, every age has its own perfect fidelity»(Sterne, 2003)[s.222]

Den perfekte lyden ligger da hverken i lagringsmediet eller i avspillingsutstyret, men kan sies å ligge i samarbeidet som oppstår mellom sender og avsender.

5.2 Hva er fordelene med høyere samplerate?

Det er vanskelig å kategorisere ut enkeltstående kvaliteter som gir håndfaste beviser for hvorfor en høy samplerate er gunstig. Det kan virke som om det er kultivert frem en idé om at et binært system trives bedre hvis det har flere holdepunkter, men dette er mine egne tanker om temaet. Det er gjort mange undersøkelser og blitt forsket mye på dette teamet, og mange forfattere av lærebøker gir sine synspunkter på saken. Jason Corey skriver i sin bok *Technical Ear Training* om høyere samplertes:

There have been a number of heated debates about the advantages or benefits of high sampling rates in digital audio. The compact disc digital audio format specifies a sampling rate of 44.100 kHz and a bit depth of 16 bits per sample, according to the Red Book CD standard. As recording technology has evolved, it has allowed recording and distribution of audio to listeners at much higher sampling rates. There is no question that bit depths greater than 16 bits per sample improve audio quality, and engineers typically record with at least 24 bits per sample. As an exercise, compare a 24-bit recording to a dithered down 16-bit version¹⁸ of the same recording and note audible differences (Corey, 2010) [s.154].

Denne øvelsen som Corey foreslår, hvor man reduserer bitdybden på et spor fra 24 til 16 og sammenligner de to, er en god metode for å lære seg forskjellen mellom de to formatene. Jeg tviler heller ikke på at det finnes forskjeller mellom de to signalene, men det bør følge med en avveining om hvorvidt forskjellene er store nok til at det er verdt de ekstra ressursene. Det er fort gjort å tenke at det er like greit å sample høyere bitrate fordi datakraft er rimelig, men denne tankegangen er lite logisk hvis det ikke er tungveiende grunner for å velge større bitdybde. For veldig mange musikkstykker og lyttesituasjoner, vil ikke de ekstra bitsene ha noen effekt, de vil bare gjøre filene større og tyngre for datamaskinen å behandle. Jeg tenker her på musikkstykker med et normalt frekvensinnhold hvor små nyanser ikke er i fokus, og jeg tenker på lyttesituasjoner utenfor et optimalt lytterom.

Selv hadde jeg denne boken som faggrunnlag da jeg studerte miksing og mastring, og jeg har erfaring med alle øvelsene i boken. Corey's synspunkter veier tungt for meg, men samtidig er jeg skeptisk til å legge for mye vekt på oppløsning når det kommer

¹⁸Corey bruker her begrepet «dithered down» for å beskrive et ledd i bitredueringen. Det er en henvisning til en prosess hvor man tilfører et lavnivå, kontrollert støysignal for å frikople kvantiseringsstøy fra signalet (Halmrast, 2018)[s.375]

til lyd kvalitet. Enhver kvalitet lar seg ikke direkte oversette til bedre opplevelser, selv om det har potensiale til å gjøre det.

With any of these comparisons, it is easier to hear differences when the audio is reproduced over high-quality loudspeakers or headphones. Lower-quality reproduction devices do not allow full enjoyment of the benefits of high sampling rates. High-quality reproduction systems do not have to be expensive, especially in consumer systems.[ibid.]

Som de fleste andre gjør, påpeker også Corey her at det er lettere å høre forskjeller over et par gode høyttalere eller hodetelefoner. Gode høyttalere er normalt mer følsomme, har raskere respons og har større frekvensområde enn rimeligere høyttalere. Samtidig åpner et slikt argument opp for en logikk som tilsier at hvis du ikke hører forskjellen, har du heller ikke dyrt nok stereoanlegg. Nå skriver han jo også for at et godt anlegg ikke nødvendigvis vil være kostbart, men forrige argument kan også brukes her. I tillegg vil jeg tilføye at lytterrommet og lytteposisjon har vel så mye å si for at stereobildet skal fungere, kanskje er det enda viktigere enn et par gode høyttalere. Ved høyere samplerrates vil også høyere frekvenser kunne få mer spillerom, og alt ettersom høyttalerene evner det, vil også de høye frekvensene kunne gjenskapes i lytterrommet. Om lytterrommet ikke er riktig behandlet og optimal lytteposisjon ikke kan akkommoderes, vil det økede frekvensomfanget ikke kunne gi den ønskede dybdeillusjonen i musikken, fordi det er gjort kompromisser med høyttalerplassering og demping av uønskede reflekser.

5.2.1 Høy samplerate i studio

Det var i innspillingsfasen at høyere oppløsning på signalet først ble anvendt i større skala. Senere ble det distribuert videre til forbrukere i et forsøk på å bringe studiolyd ut til flere. Ved å bruke høyere samplerate og bitdybde fikk systemet mer spillerom, og dette kunne fungere som et sikkerhetsnett mot uforutsette hendelser, som plutselige økninger i volum eller feil bruk av mikrofoner. Senere er høyoppløselig blitt det gjeldende formatet i mange profesjonelle studioer og det kommer også forbrukerne til gode i form av en trygghet som sier at de får flere bits inn til sin avspillingsen-

het. Hvorvidt dette objektivt er årsaken til en kvalitetsøkning er utenfor rekkevidden til denne oppgaven. Det vil kreve store mengder eksperimentering og lytteprøver, og selv ikke da vil det kunne ta høyde for individuell smak og endringer av lyttekultur, og talløse andre faktorer ¹⁹.

Det er viktig å påpeke at ikke alle er like enige i at høy samplerate faktisk har noen positiv effekt som ikke kan oppnåes ved smartere bruk av dataene og bedre komponenter i filtrene som brukes til å begrense båndbredden før analog-til-digital konvertering. Ethan Winer skriver i boken *The Audio Expert* om sitt syn på saken:

[...] Higher rates are also used, mostly by people who believe that ultrasonic frequencies are audible or that a bandwidth higher than is actually needed somehow improves the sound quality for frequencies we can hear. Common high sample rates are 88.2 kHz and 96 kHz [...] Even higher sample rates are sometimes used – 192 kHz and 384 kHz – though clearly there's no audible benefit from such overkill, and that just wastes bandwidth and hard drive space. Handling that much data throughput also makes the computer DAW²⁰ work harder, limiting the number of tracks you can work with. (Winer, 2018) [s.246]

Han skriver videre om bitdybde: «Although 16 bit offers a sufficiently low noise level ²¹ for almost any audio recording task, many people use 24 bits because, again, they believe the additional bits offer superior fidelity.[ibid.]». Han understreker at antall bits ikke har sammenheng med fideliteten, men kun signal-til-støy ratioen, som i de fleste tilfeller er bedre enn bakgrunnstøyen i ethvert studio. I situasjoner utenfor studio, hvor for eksempel et orkester skal tas opp, bruker han riktignok 24 bit, for å ha ekstra takhøyde for transienter. Et poeng til er at de fleste lydkort har et eget støygulv som er over det støygulvet en 24 bit oppløsning teoretisk har. Dermed vil den innebygde støyen allikevel trumfe over det store dynamiske omfanget en 24 bit konvertering kunne hatt. Dermed tenker jeg at en 24 bit datastrøm vil føre til et betydelig tap av data for alle de som strømmer ved sub-optimale forhold.

Det er allikevel mange musikkprodusenter som ønsker å bruke høyere samplerate, og sier de hører en tydelig forskjell. Eric 'Mixerman' Sarafin er en av dem, og han har

¹⁹Fraunhofer gruppen arbeidet i mange år, og tett med de største aktørene i musikkbransjen, med lytteprøver og kontrollgrupper for å optimalisere MP3-formatet (Sterne, 2012)

²⁰DAW, forkortelse for Digital Audio Workstation

²¹Ethan Winer skriver ofte om bitdybde som et mål på dynamisk omfang og en måte å angi signal-til-støy ratioen.

også i sin bok *Zen, and the Art of Mixing* understreket flere ganger hvor viktig det er å være i flyten når han jobber med miksing, og alt som kan få jobben til å gå raskere og mer effektivt må prioriteres. I boken inkluderer det alt fra å være en likandes kar til å kjenne utstyret en jobber med, og da er det bedre å velge høyere samplerate, om enn ikke for mer enn å bare være på den sikre siden. I et innspillingsstudio vil den største oppgaven alltid være å registrere og lagre musikk, som oftest i en så god kvalitet som mulig. Det er mange oppfatninger om hva som representerer en god kvalitet best når det kommer til valg av innspillingsformat. Det er en trend nå for å sample i så høy oppløsning systemet tillater, om enn for ikke for annet enn å være på den sikre siden. I ett innspillingsstudio eller mastringsstudio vil også forutsetningene for å høre forskjellene være bedre også, i tillegg kan man havne i en hypersensitiv tilstand, der små detaljer blir tydeligere på bekostning av det store bildet. Produsent Erix 'Mixerman' Sarafin har vært en utadgående stemme innen lydproduksjon, med en bokserie og en rekke podcaster, ved siden av en betydelig arbeidskatalog. I 2017 skriver han blant annet om samplertes i artikkelen, «But How Do You Feel?»:

[...] Further, there's clearly an audible difference between 48k fidelity and 44.1k, which will be readily discernible by most top professionals. [...] But anyone who has trained their ears over time and who has the physical and mental abilities to be good at evaluating sound, can readily hear the difference between 44.1k, 48k, 16bit and 24bit in a proper critical listening environment, on music they are super familiar with, particularly when that professional has a stake in the outcome of that music(Sarafin, 2017)

Dette utdraget er hentet fra en artikkel på Sarafin's hjemmeside, og hele artikkelen er et svar til en annen personlighet fra det profesjonelle lydmiljøet, Ethan Winer. De to hadde en lang og opphetet diskusjon om lydteknikk som strakk seg over mange år. Grunnen til at jeg velger å ta med dette sitatet er at Sarafin kommer med noen svært viktige poenger. Han skriver at det er spesielt viktig å være sikker på at formatet er riktig, hvis man har investert i resultatet. Som oftest vil det si at man velger en høyere samplertes for å være på den sikre siden. Et annet viktig poeng han gjør, er at i de rette omgivelsene er det mer åpenbart hva som er den beste kvaliteten. I tillegg må man være godt trent for oppgaven og materialet må være godt kjent og familiært. I en

profesjonell sammenheng vil det si et prosjekt man jobber tett med, og høy samplerate er ikke en garanti for at det lar seg oversette til bedre kvalitet utenfor studioet.

Ethan Winer, som Sarafin diskuterte med, har også skrevet bøker og jobbet med lyd hele livet. Jeg vil komme tilbake til han i et senere kapittel, men for nå vil jeg si at han også er anerkjent, kanskje mer som akustiker og bygger av studioutstyr, enn som musikkprodusent. at det blir så mye friksjon mellom fraksjoner av et lydmiljø, skyldes ofte at det er emosjonelle og finansielle investeringer som er gjort og som må beskyttes. Sarafin har skrevet i sin bok mye om den emosjonelle siden av musikk og viktigheten av det følelsesmessige anslaget en sang må ha for å være suksessfull. Selv om en stor del av boken også handler om tekniske aspekter ved musikkproduksjon, var det ikke på et teknisk grunnlag han kunne føre en diskusjon mot Winer. Winer har en mye dypere teknisk forståelse og han fokuserer mer på alt det som kan måles og her ligger kjernen i mange harde debatter. For ordens skyld refererer jeg også til Winer's svar til Sarafin, som også er det siste utspillet i denne debatten, såvidt meg bekjent(Winer, u.d.-b).

5.2.2 Diminishing returns, lavere avkastning

Flere eksempler på forfattere som advokerer for høyere samplerate er Bobby Owsinski i boken *The Recording Engineers Handbook*, som er en lærebok i en serie som omhandler de fleste aspekter ved studioarbeid, fra innspillingsfasen, via miksefasen og til mastringsfasen og klargjøring for distribusjon. Det er et anerkjent og veletablert volum, og som blir brukt i både undervisning og selvstudier. Når det kommer til samplertes, føler jeg at han følger trenden med å anta at antall biter man kan dele et signal opp i, jo bedre blir lyden. Jeg vil kritisere og si at selv om lydbølgen er bedre representert betyr ikke det at lyden automatisk blir bedre. Dette er et subjektivt kriterium og vil det alltid være åpent for diskusjon. Owsinski skriver:

TIP: The higher the sampling rate, the better representation of the analog signal and the greater the audio bandwidth will be, which means it sounds better!(Owsinski, 2014)[s.70]

Owsinski skriver også at kvalitetsøkningen av høyere samplerate avhenger også

av høyttalere, lyttemiljø, signalgang, og til og med hvilken type musikk som spilles. I tillegg krever høyere oppløsning kraftigere datamaskiner, noen programvarer er ikke beregnet på det, og det er færre lydspor som datamaskinen klarer å håndtere, enn om en hadde valgt en lavere oppløsning. Allikevel har 96 kHz blitt den nye bransjestandarden, fordi strømmetjenester som iTunes oppfordrer nå til å benytte høyoppløselige filer [ibid.].

I dette ligger det også et ankepunkt. Ved alle store overganger mellom formater har det også vært mye større forskjeller enn det er mellom de forskjellige digitale formatene. Ta for eksempel overgangen fra grammofon til digital, som var en mye større forandring og fremskritt, eller overgangen fra grammofon til LP, hvor nye materialer og produksjonsmetoder ga lengre levetid, bedre signal-til-støy ratio, mulighet for stereoinnspillinger (senere også kvadrofoni), for å nevne noen²². Alle kvalitetsøkningene som økt samplerate har, er avhengig av ytre faktorer, som lyttemiljø og avspillingsutstyr, men da kan det gjøres ett poeng av at forskjellene kan også ligge i måten mastringen blir utført, og ikke i selve oppløsningen. Det er ihvertfall en god idé, å ha flere kriterier enn samplerate og bitdybde når lyd kvalitet skal avgjøres.

5.3 Hva er fordelene med lavere samplerate?

Den mest åpenbare fordelen ved å benytte en lavere samplerate er at størrelsen på filene holdes små. Dette er gunstig både for å utnytte lagringsplassen og for at en datamaskin lettere skal kunne jobbe med lydfile. I figur 13 som viser bitraten til forskjellige formater er det lett å se hvordan filstørrelsen kan vokse ut av proporsjon. Selv om data over internett er blitt mer tilgjengelig er det fortsatt en endelig ressurs som bør forvaltes forsvarlig.

Som jeg beskriver i kapitlet om *Båndbredde og lagring* 2.6 og *Jevon's effekt*, er det grunn for å bruke de midlene en kan for å begrense bruken av båndbredde. En måte å gjøre dette på er ved å begrense størrelsen på filene en sender og mottar. Dette kan

²²I tillegg veide platene mindre, var tynnere, hadde mer spilletid per side, og således kunne de enklere distribueres, lagre og transporteres

Stereo

Formula:

Bits per sample x samples per second = bits per second x 2 channels = bits per second of stereo / 8 = Bytes per second of stereo x 60 seconds = Bytes per minute of stereo x 60 minutes = Bytes per hour of stereo.

Settings	Bitrate	File size per second	File size per minute	File size per hour
16 bit, 44.1 KHz	1,411.2 Kbps	176.4 KB	10.584 MB	635.04 MB
16 bit, 48 KHz	1,536 Kbps	192 KB	11.520 MB	691.2 MB
24 bit, 48KHz	2,304 Kbps	288 KB	17.28 MB	1.036 GB
24 bit, 96KHz	4,608 Kbps	576 KB	34.56 MB	2.0736 GB

Figur 13: Oversikt over filstørrelser i stereo PCM-format(AudioMountain.com, 2021)

gjøres ved å velge lavere oppløsning på musikken en strømmer og lagrer.

Et annet aspekt som jeg mener er viktig er å se på utviklingshistorien for lydteknologi når en velger samplerate. Det er mange andre elementer som er viktigere enn høy samplerate, slik som

Det er heller ingen grunn for å velge multipler av 48 kHz over multipler av 44.1 kHz på grunnlag av at 48 er et rundere tall. Det er en myte at datamaskiner kan regne lettere med dette tallet, det kan snarere være motsatt på grunn av alle primtallsfaktorene som går inn i 44.100. Jeg fant ingen spesifikke tester på dette, men som Bob Katz påpeker i kapitlet *How good is 44.1 kHz sample rate: The answer: A lot better than I used to think*, det er mye mer som avgjør lyd kvalitet enn samplerate. I stedet for å snakke om bedre eller dårligere samplerate, kan en heller snakke om kvaliteten på komponentene rundt.

Can 44.1 kHz ever sound equal to 96 kHz? It may be impossible to find out without building a custom, discrete DAC. A more effective question would be: Is 192 kHz necessary?(Katz, 2007)[s,253]

Dette sitatet er fra 2007 og det er godt mulig at det er gått ut på dato og at tendensen har skiftet mot høyere samplertes nå. Når han i tillegg blander sammen samplertes i forskjellige multipler (44.1 og 48), så tyder dette også på at det ikke er i sampleraten de virkelig avgjørende faktorene for lyd kvalitet ligger. Det er av større betydning at systemet greier å benytte alle dataene effektivt.

6 Konklusjon

I denne masteroppgaven skulle jeg svare på om det kunne finnes en optimal samplerate og eventuelt da hva den ville være. Dette var et ambisiøst foretakende, og det viste seg at dette er et bevegelig mål, og kun kan besvares med: -«finnes det en optimal samplerate i forhold til det problemet du søker å løse?».

Jeg oppdaget at det var mange forskjellige oppfatninger av hva som var viktig, og det lykkedes meg ikke å finne en samplerate som kunne dekke alle behov. Problemet er at det er for mange variabler, og det bare innen de kriteriene jeg har berørt. Det eksisterer mange flere som jeg ikke har berørt.

Jeg har forsøkt å trekke noen historiske tråder, og jeg oppdaget tidlig noe som jeg fra før hadde en mistanke om. Nemlig at idéen og tanken om en perfekt reproduksjon har alltid fulgt sammen med teknologien som brukes til lydproduksjon. Helt siden fonografen ble oppfunnet på 1870-tallet har alle produsenter, fra Edison til Sony, forsøkt å overbevise sine kunder om at deres produkter er autentiske og umulige å skille fra originalen.

Slike påstander om sann fidelitet reiste igjen visse logiske brister. Om det er mulig å lage perfekte kopier, hva er da originalen? Dette brakte spørsmålet videre til perfekte kopier i det digitale domenet kan regnes som originaler. Her forsøkte jeg å vise at selv diskrete pulser er bygget opp av kontinuerlige signaler og at overgangen mellom analog og digital er glidende.

Så konklusjonen min blir at mitt forskningsspørsmål ikke lar seg besvare i en enkel oppgave som dette.

6.1 Lyd kan ikke lagres

Jeg har også forsøkt å illustrere at lyd kun kan eksistere som et momentant fenomen, med andre ord, lydbølgen må være i bevegelse for at vi skal kunne oppfatte den. Det er ikke mulig å lytte til en datafil uten at luften foran høyttalerene settes i bevegelse igjen, og lyden igjen er akustisk.

Det vi lagrer er kun en representasjon av lyden, og lagringen kan enten være analog eller digital, og her ligger også forskjellene mellom forskjellige medier. Alt som skjer på vei til lagringen er også likt, uansett om lagringsmediet analogt eller digitalt.

Hver nye lagringsmåte brakte også med seg nye metoder for å etterarbeide eller manipulere lydsignalene på, men hver nye lagringsmåte innebar også en abstraksjon av et fysisk element som det foregående lagringsmediet var avhengig av. På denne måten mente jeg at vi ikke kunne nærme oss en bedre fidelitet på et område uten at vi samtidig fjernet oss vekk fra den samme fideliteten på et annet område.

Digital teknologi har fjernet seg lengst fra det fysiske signalet og erstattet det med en representasjon bestående av tall. Dette har gitt oss en lydgjengivelse med en uovertruffen frekvensrespons og signal-til-støy ratio, men likevel mener jeg at vi ikke har kommet nærmere en perfekt reproduksjon av den opprinnelige lyden, mye fordi vi som lyttere også har utviklet oss i samme takt som teknologien rundt oss. Vi må «spille på lag» med teknologien, lydig sette oss i «the sweet spot» og gang på gang «la oss begeistre» for at illusjonen skal kunne opprettholdes.

Etter at et signal er digitalisert, går det inn i en tilstand av lagringer. Så lenge signalet er representert som en serie med tall, er det også å regne som lagret. Det vil si at representasjonen av signalet går fra en tilstand til en annen i diskrete steg, uten noen mellomverdier slik som i en kontinuerlig tilstand. Dette har den fordel å ha et «signal» helt koplet fri fra fysiske begrensninger som treghet, masse og friksjon helt til signalet skal gjenopprettes igjen fra dataene. I denne tilstanden kan også bearbeidinger skje ved hjelp av regnestykker og beregninger uten at de opprinnelige dataene forandres, forandringene skjer bare i metadata, og således holdes representasjonen intakt. Videre kan en digital representasjon av et signal kopieres, sendes, komprimeres og lagres veldig effektivt, så det er mange fordeler ved å digitalisere lyden.

6.2 Avsluttende refleksjoner vedrørende egen oppgave

Nå som jeg anser meg som ferdig med denne oppgaven innser jeg at jeg egentlig bare har berørt en ørliten del av hva som definerer samplerate. Jeg har muligens gått litt for mye opp i eksotiske tolkninger av hva et signal kan være, og ikke satt av nok plass til å definere hva samplerate av det signalet faktisk er i en arbeidssammenheng.

Videre har det også vært mange kilder å forholde seg til og jeg tror oppgaven vil bære preg av dette. Jeg har lest mange bøker fra perm til perm for denne oppgaven og jeg kunne med fordel ha strukturert lesingen bedre med hensyn på notering underveis. Pensumlitteraturen har jeg valgt ut fra bøker jeg kjenner fra studier, og videre har jeg fulgt de forfatterene og forskerene som har vært henvist til i sitater og utdrag. Jeg kunne også ha gjort flere søk på nøkkelord fra oppgaven, men jeg følte at jeg hadde nok materiale til å få en så veloverveid konklusjon som jeg greide til slutt.

Det var også hensikten å bringe inn noe mer matematikk i oppgaven, da jeg har dette som støtteemne (40 gruppe) i bachelorgraden min. Grunnen til at jeg valgte bort dette er at det er en fare for at blir en annen type oppgave og jeg mente at jeg ikke hadde nok bakgrunn for dette.

7 Kildeliste

Kildeliste

Aliasing. (2021). <https://en.wikipedia.org/wiki/Aliasing>. (Hentet 27-5-23:35)

Anderson, T.J. (2015). *Living stereo*. I (s. 107–124). Bloomsbury Academic. (An imprint of Bloomsbury Publishing Inc)

AudioMountain.com. (2021). Audio file size calculations. <https://www.audiomountain.com/tech/audio-file-size.html>. Forfatter. (Hentet 8-6-2021)

Audio myths and daw wars. (2021). https://www.image-line.com/fl-studio-learning/fl-studio-online-manual/html/app_audio.htm. (hentet 24-5-2021 fra manualen til FL-Studio 20)

Corey, J. (2010). Audio production and critical listening (1. utg.). Burlington, US: Focal Press. (Focal Press is an imprint of Elsevier inc.)

Devine, K. (2019). Decomposed. Massachusetts: Massachusetts Institute of Technology.

Emerick, G. & Massey, H. (2002). Here, there and everywhere. New York: Avery. (an imprint of Penguin Random House LLC)

Ernst, W. (2016). Chronopoetics. New York: Rowman and Littlefield International Ltd.

Farmelo, A. (2012). The problem with a-b'ing and why neil young is right about sound quality. <https://tapeop.com/blog/2021/05/07/problem-bing-and-why-neil-young-right-about-sound/>. Tape Op Magazine. (Hentet 7-6-2021 22:27)

-
- Feynman, R. (1985). Idiosyncratic thinking. <https://www.youtube.com/watch?v=EKWGGDXe5MA>.
- Firstsounds. (u.d.). The phonoautograms of Édouard-léon scott de martinville. <http://www.firstsounds.org/sounds/scott.php>. (hentet 3-5-2021, 22:11)
- Grajeda, T. (2015). *Living stereo*. I (s. 37–63). Bloomsbury Academic. (An imprint of Bloomsbury Publishing Inc)
- Halmrast, T. (2018). *Klangen [programvarehåndbok]*. (Kompendium for *Lydteori og Lydlære 1 og 2* ved Inst. for Mus.Vit, UIO)
- Hjelm, M. & Haave, B. (1978). Stereo hi-fi. Stabekk: NKI-forlaget. (Originaltittel: *Master Hi.Fi Installations*; G.J.King; Butterworth & Co. (Publishers) Ltd., 1976)
- Holman, T. (2008). Surround sound (2. utg.). Burlington(US) AND Abingdon(UK): Taylor & Francis. (Focal Press is an imprint of the Taylor & Francis Group, an informa business)
- Holman, T. (2010). Sound for film and television (3. utg.). Burlington(US) AND Abingdon(UK): Taylor & Francis. (Focal Press is an imprint of the Taylor & Francis Group, an informa business)
- Johansen, K. (2021). Datalagring som evighetsprosjekt. Aftenposten Innsikt, 14(4), 104–112.
- Katz, B. (2007). Mastering audio. Burlington, MA, USA: Focal Press. (Focal Press is an imprint of Elsevier)
- Lehrmann, P.D. (2008). The emperor's new sampling rate. <https://www.mixonline.com/recording/emperors-new-sampling-rate-365968>. Mixonline. (Hentet 24-5-2021 18:00)
- Mihay. (2021). Square wave generator with 7400 series chip. <http://www.antimath.info/electro/square-wave-generator-with-7400-series-chip>. AntiMath thoughts,simple. (Hentet 26-5-2021 18:48)

- Montgomery, C. (u.d.). D/a and a/d | digital show and tell. <https://www.youtube.com/watch?v=cIQ9IXSUzuM>. FL Studio by Image-Line Software. (Hentet fra Youtube.com 24-5-2021 23:07)
- Moore, C.W. (1959). Mountain-to-mohammed. High Fidelity and Audiocraft, 9(1), 18–20. <https://worldradiohistory.com/Archive-All-Audio/Archive-High-Fidelity/50s/High-Fidelity-1959-Jan.pdf>. (Hentet 3-5-2021 18:38)
- Owsinski, B. (2014). The recording engineer's handbook (3. utg.). Boston: Cengage Learning.
- Pohlmann, K.C. (2000). Principles of digital audio (4. utg.). New York: McGraw-Hill.
- Rate. (2018). <https://snl.no/rate>. Store norske leksikon. (Hentet 28.april 2021)
- Rothenbuhler, E.W. & Peters, J.D. (1997). Defining phonography. The Musical quarterly, 81(2), 242–243. <https://academic.oup.com/mq/article/81/2/242/1062261>. (Downloaded by University of Oslo Library user on 04 June 2021)
- Sampling. (2018). <https://snl.no/sampling>. Store norske leksikon. (Hentet 28.april 2021, skrevet av Ragnar Johnsen)
- Sandberg, M.B. (1991). Missing persons: spectacle and narrative in late nineteenth-century scandinavia. University of California, Berkeley. (Referanse fra Rothenbuhler & Peters; Defining Phonography: An Experiment in Theory)
- Sarafin, E. (2017). But how do you feel. <https://mixerman.net/but-how-do-you-feel/>. (Wieved 2021-01-07, 01:57)
- Simonsen, T. (2012). Vad är egentligen ett fonogram? I G. Ternhaug & J. Wingstedt (red.), På tal om musikkproduktion (s. 107–122). Göteborg: Bo Ejeby Förlag.
- Simonsen, T. (2021). Lydspor. <https://snl.no/grammofon>. Store Norske Leksikon. (Hentet 19. april 2021, Lisens CC BY SA 3.0)

Sterne, J. (2003). The audible past. North Carolina: Duke University Press Durham and London. www.dukeupress.edu.

Sterne, J. (2012). The meaning of a format. North Carolina: Duke University Press Durham and London. www.dukeupress.edu.

unknown. (u.d.). The edison phonograph. <https://www.houzz.com/products/the-edison-phonograph-poster-print-by-unknown-18x24-prvw-vr~123553208>. (hentet 3-5-2012, 21:44)

Wadell, G. (2013). Complete audio mastering (3. utg.). New York: McGraw-Hill Education.

Weisstein, E.W. (2021). Fourier series – square wave. <https://mathworld.wolfram.com/FourierSeriesSquareWave.html>. Mathworld–A Wolfram Web Resource. (Hentet 26-5-2021 18:42)

What are the types of signals. (2021). <https://www.quora.com/What-are-the-types-of-signals>. (Hentet 27-5-2021 19:33)

Winer, E. (u.d.-a). Does dither really matter? <https://ethanwiner.com/dither.html>. (Hentet 27-5-2012 22:13)

Winer, E. (u.d.-b). Mixer man exposed. <http://ethanwiner.com/mixer man.htm>. www.ethanwiner.com. (Wieved 2021-01-07, 01:57)

Winer, E. (2018). The audio expert (2. utg.). New York: Taylor & Francis. (First edition published by Focal Press 2012)

Appendix

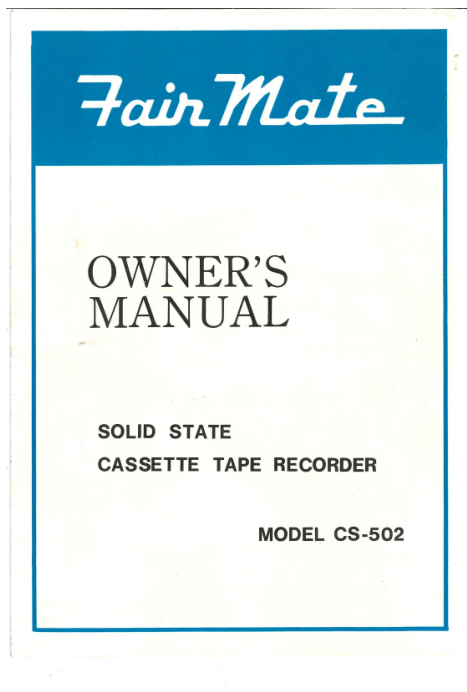
8 Tillegg



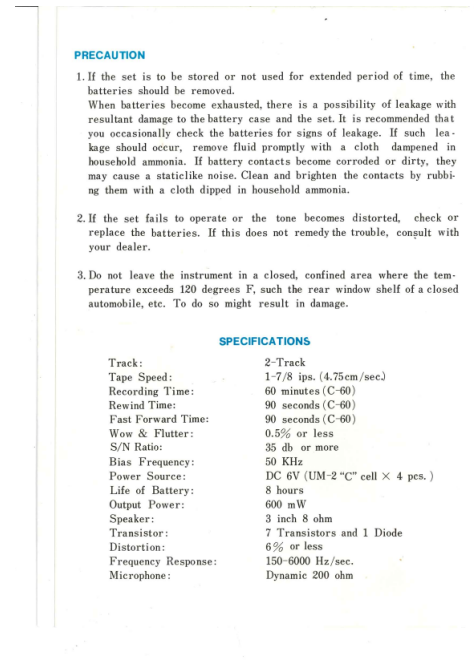
(a) Scannet fra instruksjonshefte/manual

(b) Scannet fra datasheet/manual.

Figur 14: Frekvensområde 40 Hz-16 kHz og 46 dB dynamisk spenn.



(a) Scannet fra instruksjonshefte/manual



(b) Scannet fra datasheet/manual.

Figur 15: Frekvensområde 150-6000 Hz og 35 dB dynamisk spenn.

Skrevet i L^AT_EX