

MASTER

Het filteren van pianosignalen m.b.v. een digitale computer

Scha, R.J.H.

Award date:
1970

[Link to publication](#)

Disclaimer

This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Afdeling der Elektrotechniek
Technische Hogeschool
Eindhoven

GROEP METEN EN REDELEN

Het filteren van pianosignalen
m.b.v. een digitale computer.

R.J.H. Scha

Rapport van het afstudeerwerk
uitgevoerd van maart 1969 tot mei 1970
in opdracht van prof.dr.ir. P. Eykhoff
onder leiding van ir. A.A. van Rede.

Samenvatting.

Aan de hand van de eigenschappen van pianosignalen, zoals die uit een literatuurstudie naar voren komen, worden enkele mogelijke filtermethodes voor deze signalen beschouwd. De methode die daarvoor het efficiëntst leek, werd als basis-gebruikt voor een rekenmachineprogramma. De werking van dit programma, en van de conversie-programma's voor input en output, wordt geschetst.

Inhoudsopgave.

1. Inleiding.	1
2. De beschrijving van de pianotoon.	2
2.1. Inleiding.	2
2.2. De grondfrequentie.	2
2.3. De frequenties van de boventonen.	3
2.4. De tijdsafhankelijkheid van de amplitudes.	3
2.5. De amplitudes van de verschillende spectrale componenten.	3
2.6. De faschoeken van de boventonen.	4
3. Detectiemethoden.	5
3.1. Inleiding.	5
3.2. Matched filter toegepast op de pianotoon.	5
3.3. Matched filter toegepast op de afzonderlijke componenten.	6
3.4. Fourier-transformatie.	7
3.5. Additieve correlatie.	7
3.6. Conclusie.	8
4. Synthesemethoden.	9
5. Schattingen, benodigd voor de synthese.	10
5.1. Schatting van de grondfrequentie.	10
5.2. Schatting van de golfvorm.	10
5.3. Schatting van het amplitudeverloop.	10
6. Het programma filter2.	11
7. De analoog-digitaal-conversie.	13
8. De digitaal-analoog-conversie.	13
9. Het gebruik van de software van het hybride systeem in PL/I.	14
9.1. Inleiding.	14
9.2. Arithmetische parameters van de subroutines.	14
9.3. Input- en output- arrays van de subroutines.	14
10. Conclusies.	16
11. Literatuur.	17

1. Inleiding.

Oude grammofoonplaten hebben hinderlijke ruis. Gevraagd wordt een filtermethode voor pianomuziek op deze grammofoonplaten.

De piano heeft een aantal snaargroepen, die op willekeurige momenten met een hamer meer of minder hard in beroering kunnen worden gebracht. Dit betekent dat er 88 signalen gestart kunnen worden op willekeurige tijdstippen en met willekeurige vermogens. Voorts zijn er een aantal van deze snaargroepen van dempers voorzien, die het vermogen zeer snel tot nul kunnen reduceren. Dit kan eveneens op een willekeurig tijdstip gebeuren.

Het filter zal dus de volgende functies moeten vervullen:

- a. een analyse van de 88 tonen.
- b. schatting van de starttijdstippen, vermogens en eindtijdstippen.
- c. een hernieuwde samenstelling van het signaal met weglating van de ruis.

De eigenlijke signaalverwerking vindt plaats in de digitale computer, d.m.v. de PI/I-procedure filter2. De inputgegevens van deze procedure staan op een digitale band. Deze gegevens zijn de samplewaarden van het te verwerken signaal. De outputgegevens van het programma worden eveneens op een digitale band gebracht. Deze opslagmethode is gekozen vanwege het zeer grote aantal data dat moet worden bewaard.

De eigenlijke signaalverwerking wordt dus nog voorafgegaan door de vervaardiging van de digitale inputband, en gevolgd door de digitaal-analoog-conversie vanaf de outputband. Moeilijkheden in verband hiermee hadden tot gevolg, dat onze filtermethode niet in de praktijk beproefd kon worden. De AD- en DA-conversie worden behandeld in resp. de hoofdstukken 7 en 8.

De procedure filter2 komt in hoofdstuk 6 aan de orde.

Teneinde te kunnen beslissen over de toe te passen methodes van analyse en syntese (hoofdstuk 3,4,5) is allereerst nagegaan op welke wijze de tonen van de piano kunnen worden beschreven. De piano die de signalen produceerde voor de te filteren grammofoonplaat stond niet tot onze beschikking. De specifieke eigenschappen van zijn signalen zijn dus onbekend. Om een inzicht te krijgen in pianotonen in het algemeen is een literatuurstudie gedaan. (lit.1 t/m 4) Enige daaraan ontleende gegevens kan men vinden in hoofdstuk 2.

2. De beschrijving van de pianotoon.

2.1. Inleiding.

De pianotoon kan, op grond van zijn ontstaanswijze, verwacht worden te bestaan uit de som van een aantal sinusvormige componenten, waarvan de amplituden in de tijd variëren. In complexe notatie kunnen we dit formuleren als:

$$S(t) = \sum_{k=1}^K a(k,t) \exp(i(\omega_k t + \varphi_k))$$

In de navolgende paragrafen zullen we nagaan, of we iets meer kunnen zeggen over de in deze formule voorkomende parameters.

2.2. De grondfrequentie.

De hoogte van een toon wordt bepaald door de grondfrequentie van de trillende snaren. Deze frequentie is gedefinieerd met een onnauwkeurigheid van 2 cent^{§)}: de grondfrequenties van 2 snaren die tot de zelfde toets behoren, verschillen i.h.a. niet meer dan 2 cent, en nooit meer dan 4 cent. (lit. 2)

De grondfrequentie van een toon verandert nauwelijks tijdens het spelen. Dit bij blijkt uit lit. 2, fig. 2b: de stemming van een vleugel voor en na een piano-koncert door mevr. prof. Elly Ney. De verschillen zijn meestal kleiner dan 1 cent. Voor een stuk van slechts enkele minuten zullen de verschuivingen dus verwaarloosbaar klein zijn.

Volgens de bestaande internationale konventies worden de grondfrequenties van de verschillende tonen op de volgende wijze gedefinieerd:

1. De frequentie van A_4 is 440 Hz.
2. De frequentieverhouding tussen de tonen wordt bepaald door de gelijkzwevende stemming. (Dit betekent dat tussen de 12 opeenvolgende tonen in een oktaaf telkens een frequentieverhouding bestaat.)

Op deze manier wordt de frequentie van toon nr n : $22,5 \cdot 2^{(n-1)/12}$
Om te kunnen beoordelen in hoeverre men aan deze formule a priori-verwachtingen bij de detektie kan ontleen, dient men te weten in hoeverre pianostemmers zich aan bovengenoemde konventies plegen te houden. Uit lit. 2 blijkt, dat er bij een volgens artistieke deskundigen goed gestemde piano systematische afwijkingen t.o.v. 1. en 2. kunnen optreden: men kiest de frequenties vaak iets te hoog, en men kiest de frequentieverhouding tussen de tonen vaak iets te groot. Dit resulteert in afwijkingen van maximaal 25 cent t.o.v. de volgens 1. en 2. bepaalde frequentiewaarden.

§) 1 cent = 1 % van de afstand (op logaritmische schaal) tussen 2 opeenvolgende tonen =
= $2^{1/1200}$.

2.3. De frekwenties van de boventonen.

Elke toon bevat, behalve de grondfrequentie ω , een aantal boventonen, welke in eerste benadering beschouwd kunnen worden als harmonischen van de grondfrequentie. Wanneer men dit iets nauwkeuriger bekijkt (lit.1), dan blijken er afwijkingen te zijn, die verklaard kunnen worden uit de stijfheid van de snaar, en beschreven volgens:

$$\omega_k = k \cdot \omega \cdot \sqrt{1 + Bk^2}$$

Dit betekent dat, wanneer men de verschillen tussen de grondfrequenties van de verschillende snaren verwaarloost (zie 2.2),

$$s(t) = \sum_{k=1}^k a(k,t) \exp(ik\omega t \sqrt{1+Bk^2} + i\varphi_k)$$

B is een konstante die voor verschillende snaren verschillende waarden heeft. Op de door Fletcher c.s. (lit.4) onderzochte piano varieerde B van ca. $5 \cdot 10^{-4}$ tot $5 \cdot 10^{-5}$.

2.4. De tijdsafhankelijkheid van de amplitudes.

Bij het aanslaan van de toon treden inslingerverschijnselen op, vervolgens dempt de toon langzaam uit, tot de demper op de snaren komt. De inslingerverschijnselen, die kunnen duren tot ca. 0,07 sek., lopen bij verschillende piano's dusdanig uiteen, dat ze bij de detektie in elk geval buiten beschouwing gelaten moeten worden. (lit. 3, fig. 9, 10, 11.) Het verdere verloop ligt voor elke harmonische van elke toonhoogte binnen bepaalde grenzen, die echter zo wijd zijn, dat de onderlinge verschillen tussen piano's aanzienlijk kunnen zijn.

Voor wat betreft de syntese is het van belang te vermelden, dat Fletcher c.s. (lit.4) nauwelijks van echt te onderscheiden pianoklanken gesynthetiseerd hebben, waarbij het amplitudeverloop voor alle spektrale componenten identiek was. (De componenten werden eerst opgeteld, daarna werd het verloop van de totale amplitude geregeld.)

2.5. De amplitudes van de verschillende spektrale componenten, onmiddellijk na afloop van de inslingerverschijnselen.

Wanneer men de amplitudes van de verschillende componenten met elkaar vergelijkt, kan men stellen dat de amplitude in het algemeen voor hogere harmonischen aanzienlijk kleiner is dan voor lagere. Voor de verschillende toonhoogten kan men zeer globaal aangeven tussen welke grenzen de gemiddelde waarde van de amplitude-afname moet liggen. (lit.4, p.760-761.) Voor wat betreft de details van het amplitudeverloop als functie van de frequentie, constateert men echter grote verschillen tussen de verschillende tonen van een piano (lit.3, fig.5) en tussen verschillende piano's (lit.3, fig. 6).

2.6. De fasehoeken van de boventonen.

Door de keuze van het tijdstip waarop $t = 0$ kan men de fasehoek van de grondtoon nul maken. De fasehoeken van de harmonischen geven dan de faseverschuivingen van deze componenten t.o.v. de grondtoon weer. Over deze fasehoeken kunnen we in de literatuur niets vinden; dit hangt waarschijnlijk samen met de omstandigheid dat ze niet waarneembaar zijn voor het menselijk oor.

De fasehoeken onmiddellijk na het inslingeren worden bepaald door de manier waarop het inslingeren plaats vindt. Dat is iets, wat voor de verschillende spectrale componenten sterk uiteenloopt. Ook tussen verschillende piano's treden in dit opzicht grote verschillen op. (lit. 3, fig. 9,10,11.) Een analytische beschrijving van het inslingerverschijnsel kan verwacht worden erg ingewikkeld te zijn. De meest redelijke aanname lijkt voorlopig, dat de verschillende boventonen een willekeurig samengestelde verzameling van fasehoeken zullen vertonen.

3. Detektiemethoden.

3.1. Inleiding.

In dit hoofdstuk zullen we enkele detektiemethoden bekijken op hun bruikbaarheid voor ons probleem. Daar we zoeken naar een methode die geschikt is om gebruikt te worden in een rekenmachineprogramma dat een signaal van behoorlijke lengte kan verwerken, zullen we vooral ook bezien, hoeveel rekentijd elk der methoden vereist.

3.2. Matched filter toegepast op de pianotoon.

De matched filter methode is de optimale methode voor de detektie van één enkel deterministisch signaal in witte ruis. (lit. 5) De methode houdt in, dat men zijn beslissingen over de aanwezigheid van het te detekteren signaal neemt op grond van de waarde van de integraal over het produkt van het gezochte signaal en het beschouwde gedeelte van het onderzochte signaal.

Het probleem dat wij ons gesteld hebben is wezenlijk ingewikkelder dan het bovengenoemde: er zijn 88 verschillende tonen, en er kunnen er meerdere tegelijk voorkomen. Dit betekent dat het er niet in de eerste plaats om gaat, de signaal-ruis-verhouding zo groot mogelijk te maken, maar dat het minstens even belangrijk is om de kans op verwarring tussen twee tonen zo gering mogelijk te maken.

Laten we desondanks nagaan of we de oplossing van ons detektieprobleem kunnen zoeken in de richting van bovengenoemde methode. Essentieel voor die methode is, dat men beschikt over beschrijvingen in het tijd domein van de te detekteren signalen. In ons geval betekent dit: voldoende goede schattingen van $a(k,t)$, ω_k en φ_k voor alle componenten van alle tonen. Wat dat betreft kan het volgende opgemerkt worden:

1. De amplitudes kan men konstant veronderstellen wanneer men een voldoende kort stuk van de toon, na afloop van het inslingeren, beschouwt. (zie 2.4)
2. Van de grootte van de aldus konstant veronderstelde $a(k,t)$ kan men zeer grove schattingen maken. (zie 2.5)
3. Van ω_k kan men schattingen maken, die een bepaalde onnauwkeurigheid hebben, maar geen verwarring tussen verschillende toonhoogten behoeven te veroorzaken. (zie 2.2, 2.3)
4. Voor grotere k wordt de schatting van slechter. (zie 2.3)
5. Over φ_k valt helemaal niets te zeggen. (zie 2.6)

De eerste 4 van deze punten houden alle 4 in dat de detektie slechter wordt dan wanneer het te detekteren signaal exact bekend zou zijn; ze impliceren niet per se dat de bovenomschreven techniek volledig onbruikbaar is. Dat laatste geldt wel voor het onder 5 genoemde punt:

Stel dat men het signaal $A(t) = \sum_k a_k \sin(k\omega t + \varphi_k)$

wil detekteren m.b.v. de benadering:

$$B(t) = \sum_k a_k \sin(k\omega t + \psi_k)$$

De integraal over 1 periode is dan:

$$\int_0^{2\pi/\omega} A(t) \cdot B(t) dt = \int_0^{2\pi/\omega} \sum_k a_k (\cos \varphi_k \sin k\omega t + \sin \varphi_k \cos k\omega t) \cdot \sum_k a_k (\cos \psi_k \sin k\omega t + \sin \psi_k \cos k\omega t) dt$$

Uit de praktijk van het berekenen van Fourier-coëfficiënten zal men zich herinneren dat de meeste termen die bij de uitwerking van de integrand ontstaan, bij het integreren weer verdwijnen; de andere leiden tot de uitkomst:

$$\frac{\omega}{\pi} \sum_k a_k^2 (\cos \varphi_k \cos \psi_k + \sin \varphi_k \sin \psi_k) = \frac{\omega}{\pi} \sum_k a_k^2 \cos(\varphi_k - \psi_k)$$

Wanneer $\varphi_k = \psi_k$, dan is dit gelijk aan: $\frac{\omega}{\pi} \sum_k a_k^2$.

Wanneer er echter niets bekend is over de relaties tussen φ_k en ψ_k , weten we in het geheel niet, welke waarde tussen $\pm \frac{\omega}{\pi} \sum_k a_k^2$ we te verwachten hebben.

3.3. Matched filter toegepast op de afzonderlijke componenten van de pianotoon.

Het bezwaar tegen de toepassing van de matched filter-techniek, naar voren gebracht in 3.2, wordt ondervangen wanneer men de spektrale componenten van het signaal afzonderlijk detekteert: dan heeft men niets te maken met de fasere relaties tussen deze componenten onderling. Deze toepassing van de methode zullen we nu nader beschouwen.

We weten reeds dat de methode optimaal is voor wat betreft de diskriminatie t.o.v. witte ruis. (Bij integratie over y samples wordt de signaal-ruis-verhouding y maal zo groot.) Van belang is nu vooral, hoe de integraal zich gedraagt wanneer we een frekwentie ω_1 willen detekteren die niet in het signaal aanwezig is, als in het signaal wel de frekwentie ω_2 (die in de buurt van ω_1 kan liggen) voorkomt. Om hier enig inzicht in te krijgen, berekenen we

$$\begin{aligned} I &= \int_0^{n \cdot 2\pi/\omega} \sin \omega t \cdot \sin(1+\Delta) \omega t dt = \frac{1}{2} \int_0^{n \cdot 2\pi/\omega} (\cos \Delta \omega t - \cos(2+\Delta) \omega t) dt = \\ &= \frac{1}{2} \left(\frac{\sin \Delta n 2\pi}{\Delta \omega} - \frac{\sin(2+\Delta) n 2\pi}{(2+\Delta) \omega} \right) = \frac{1}{2} \left(\frac{\sin \Delta n 2\pi}{\Delta \omega} - \frac{\sin \Delta n 2\pi}{(2+\Delta) \omega} \right) = \\ &= \frac{\sin \Delta n 2\pi}{\Delta(2+\Delta) \omega} \end{aligned}$$

Voor kleine Δ is dit ongeveer $\sin \frac{\Delta n 2\pi}{2 \Delta \omega}$.
Deze functie is weergegeven in fig. 2.

Een theoretisch optimale signaaldetectie betekent in elk geval dat we alle harmonischen willen detekteren, en dat we een zo groot mogelijk gedeelte van het signaal in de detectie willen betrekken. Daar de tonen gemiddeld ca. 25 harmonischen hebben, de samplefrequentie 30 kHz bedraagt, en er ca 10 stappen nodig zullen zijn om te bepalen wanneer de te vermenigvuldigen signalen in fase met elkaar zijn, komt men, wanneer men b.v. een lengte van 0,3 sek beschouwt, voor de 88 tonen op totaal $88 \cdot 10000 \cdot 10 \cdot 25 =$ ca 200 miljoen berekeningen, waarbij elke berekening bestaat uit 1 vermenigvuldiging (ca 100 μ sek) en 1 optelling (ca 10 μ sek). De totale tijdsduur bij 1 detectie wordt dan ca. 6 uur.

Daar dit volstrekt onbruikbaar is, zullen we nu nagaan, welke rekentijd er gemoeid is met een zuiniger opgezette versie van deze methode.

Beschouwen we daartoe de hierboven berekende functie I. Wanneer men bij de detectie verwarring tussen 2 tonen, die in toonhoogte slechts 1 halvetoonsafstand van elkaar verschillen, wenst te voorkomen, moet in de buurt van $\Delta = 0,06$ (dit correspondeert met 1 halvetoonsafstand tussen de frequenties) de waarde van I voldoende zijn afgenomen. Uit fig. 2 blijkt nu, dat dan $1/2n \leq 0,06$ moet zijn, d.w.z. $n \geq 8$. We moeten dus de integraal berekenen over minstens 8 perioden. Stel dat we in 10 stappen de juiste fase vinden, en dat we 5 componenten detekteren. Wanneer 1 periode p samples bevat, is op die manier het aantal

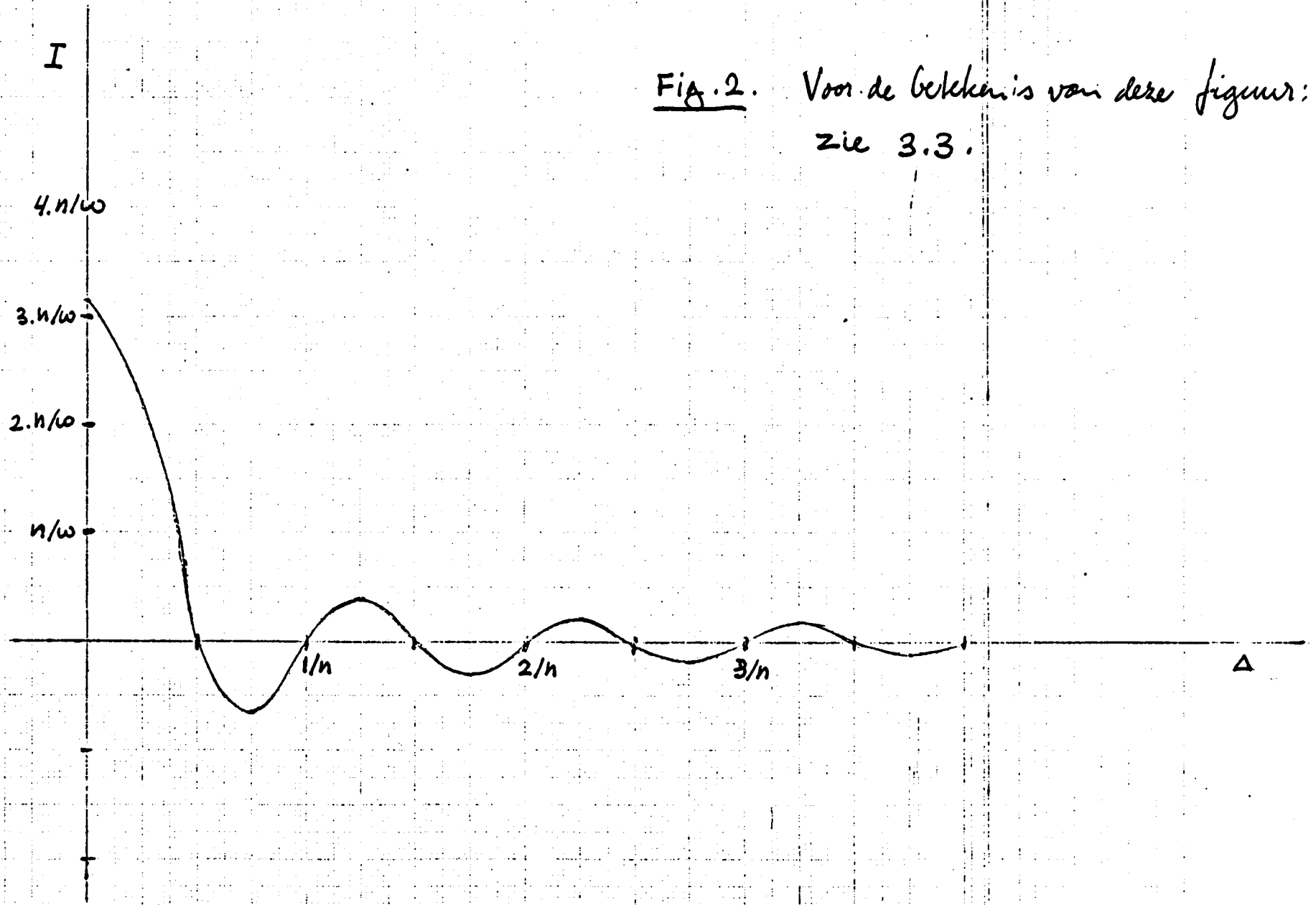


FIG. 2. Voor de betekenis van deze figuur:
zie 3.3.

bewerkingen per toon: 8. 10. 5. $p = 400 p$, waarbij elke bewerking bestaat uit 1 vermenigvuldiging en 1 optelling. Voor 88 tonen, met een gemiddelde periode-lengte van ca. 300 samples, betekent dit een rekentijd van $88 \cdot 300 \cdot 400 \cdot 110 \mu\text{sek}$ = ca 20 minuten.

3.4. Fourier-transformatie.

De op de matched filter-techniek geïnspireerde methode, uiteengezet in 3.3, komt in feite neer op het detekteren van spektrale componenten van het signaal. Men kan zich afvragen of dat niet efficiënter kan gebeuren in het frequentie-domein. We zullen daarom nu de fourier-transformatie beschouwen.

Uit N samples kan men, bij samplefrequentie f , de waarde van de fourier-getransformeerde berekenen voor $f/N, 2f/N, 3f/N, \dots, (N-1)f/N, f$. Wanneer we, bij de detektie van f_1 , onderscheid willen kunnen maken tussen tonen met opeenvolgende toonhoogten, dan moet in elk geval f/N zo klein zijn, dat f_1 en $1,06 \cdot f_1$ evenveel of meer van elkaar verschillen als de frequenties $f/N, 2f/N, \dots$. Dit betekent dat $f/N \leq 0,06 \cdot f_1$
 $N \geq f/f_1 \cdot 1/0,06$
 $N \geq 16 \cdot f/f_1$.

f/f_1 is de periodelengte van de te detekteren frequentie. Het aantal samples wat we in de transformatie betrekken moet dus minstens 16 maal zo groot zijn als deze periodelengte. Dit kost bij de Fast Fourier Transform:

$16 \cdot p \cdot \log(16 \cdot p)$ bewerkingen, waarbij onder 1 bewerking opnieuw wordt verstaan: 1 vermenigvuldiging + 1 optelling. (lit.6)

Wanneer we alleen fourier-transformeren, hebben we aan de signaal-ruis-verhouding helemaal niets gedaan. Willen we dat wel doen, dan moeten we het verogenspektrum bepalen, d.w.z. de fourier-getransformeerde van de autocorrelatie-functie. Het aantal bewerkingen wordt daardoor minstens tweemaal zo groot. Om tussen alle grondfrequenties van de piano te kunnen discrimineren, moeten we minstens 16 periodes van de laagste frequentie transformeren. De schatting van de rekentijd wordt dan: $2 \cdot 16 \cdot 1300 \cdot 2 \log(16 \cdot 1300) \cdot 110 \mu\text{sek} = \text{ca. } 3 \text{ minuten}$.

3.5. Additieve correlatie.

Wanneer een signaal bestaat uit een bepaalde frequentie en harmonischen daarvan, dan komt dit in het tijd-domein tot uitdrukking in de periodiciteit van dat signaal. Deze periodiciteit kan men gebruiken voor de detektie van het signaal. De pianotoon lijkt voldoende op het zojuist aangeduide type signaal, om ons te doen besluiten de mogelijkheden van deze methode na te gaan.

In de functie $S(t)$ kan men een periodiciteit met periode T detekteren door te berekenen de functie $X_T(t) = \sum_{n=0}^{N-1} S(t+nT)$.

Laten we dit direkt nagaan voor de pianotoon, door van

$$S(t) = \sum_{k=1}^K a(k,t) \exp i(\omega_k t + \varphi_k)$$

een zo kort gedeelte te beschouwen dat we de amplitudes konstant kunnen veronderstellen: $a(k,t) = a_k$. Verder brengen we de fasehoek onder in de complexe amplitude A_k : $a_k \exp(i\varphi_k) = A_k$

Dan is
$$S(t) = \sum_{k=1}^K A_k \exp i\omega_k t.$$

$$\begin{aligned}
 x_T(t) &= \sum_{n=0}^{N-1} \sum_{k=1}^K A_k \exp i\omega_k(t+nT) = \\
 &= \sum_{k=1}^K A_k \left\{ \sum_{n=0}^{N-1} \exp i\omega_k nT \right\} \exp i\omega_k t = \\
 &= A_k W_k \exp i\omega_k t, \\
 \text{met } W_k &= \sum_{n=0}^{N-1} (\exp i\omega_k T)^n = \frac{1 - \exp iN\omega_k T}{1 - \exp i\omega_k T} = \\
 &= \frac{\exp(-iN\omega_k T/2) - \exp(iN\omega_k T/2)}{\exp(-i\omega_k T/2) - \exp(i\omega_k T/2)} \cdot \frac{\exp(-i\omega_k T/2)}{\exp(-iN\omega_k T/2)} \\
 &= \frac{\sin N\omega_k T/2}{\sin \omega_k T/2} \cdot \exp(i(N-1)\omega_k T/2).
 \end{aligned}$$

$x_T(t)$ is dus een functie die dezelfde componenten bevat als $S(t)$. De amplitudes van deze componenten zijn echter elk vermenigvuldigd met een weegfactor W_k , waarvan de amplitude gelijk is aan

$$\frac{\sin N\omega_k T/2}{\sin \omega_k T/2}.$$

Wanneer men ω_k als variabele beschouwt, is dit (voor even N) een symmetrische periodieke functie met periode $2\pi/T$. Een halve periode van deze functie is π voor $N = 20$ weergegeven in fig. 3. Uit deze figuur blijkt dat signaalcomponenten met een frequentie die niet in de buurt ligt van $2\pi/T$ of een veelvoud daarvan, weggefilterd worden t.o.v. componenten waarvoor dat wel het geval is. Voor wat betreft de diskriminatie tussen verschillende frequenties levert deze methode geen wezenlijk andere resultaten als de matched filter-methode (vgl. fig. 2 en fig. 3), met dit verschil dat we nu niet 1 frequentie detecteren, maar een hele frequentie met al zijn harmonischen samen. De diskriminatie tussen ω en ω_{106} is voldoende groot bij $N = 20$. Dit vindt men geïllustreerd in fig. 4. Horizontaal is daar uitgezet: het nummer van de toonhoogte die overeenkomt met de frequentie $1/T$. Vertikaal is aangegeven: de energie van $x_T(t)$, met $N = 20$. Het signaal was een zaaigandspanning, met een frequentie die overeenkwam met toonhoogte 74. De bovenste stippellijn geeft aan: N^2 maal de signaalenergie van $S(t)$. De onderste geeft aan: N maal de signaalenergie van $S(t)$. De signaal - witte ruis - verhouding wordt door het N keer optellen een factor N verbeterd.

Voor het berekenen van de energie van de som over 20 periodes zijn 20.T optellingen nodig, plus nog T-maal 1 vermenigvuldiging + 1 optelling. Voor 88 tonen van gemiddeld 300 samples duurt dit $88 \cdot 300 \cdot (20 \cdot 10 + 110) \mu \text{ sek} =$
 $= \text{ca } 8 \text{ sek.}$

3.6. Konklusie.

Op grond van het extreme verschil in rekestijd (8 sek. per detectie, tegenover resp. 6 uur, 20 minuten en 3 minuten voor de andere methodes) meenden wij dat het ten zeerste de moeite waard zou zijn om te proberen, de detectie te doen geschieden m.b.v. de methode van additieve correlatie.

Fig. 3 W_k als functie van ω_k . (zie 3.5.)

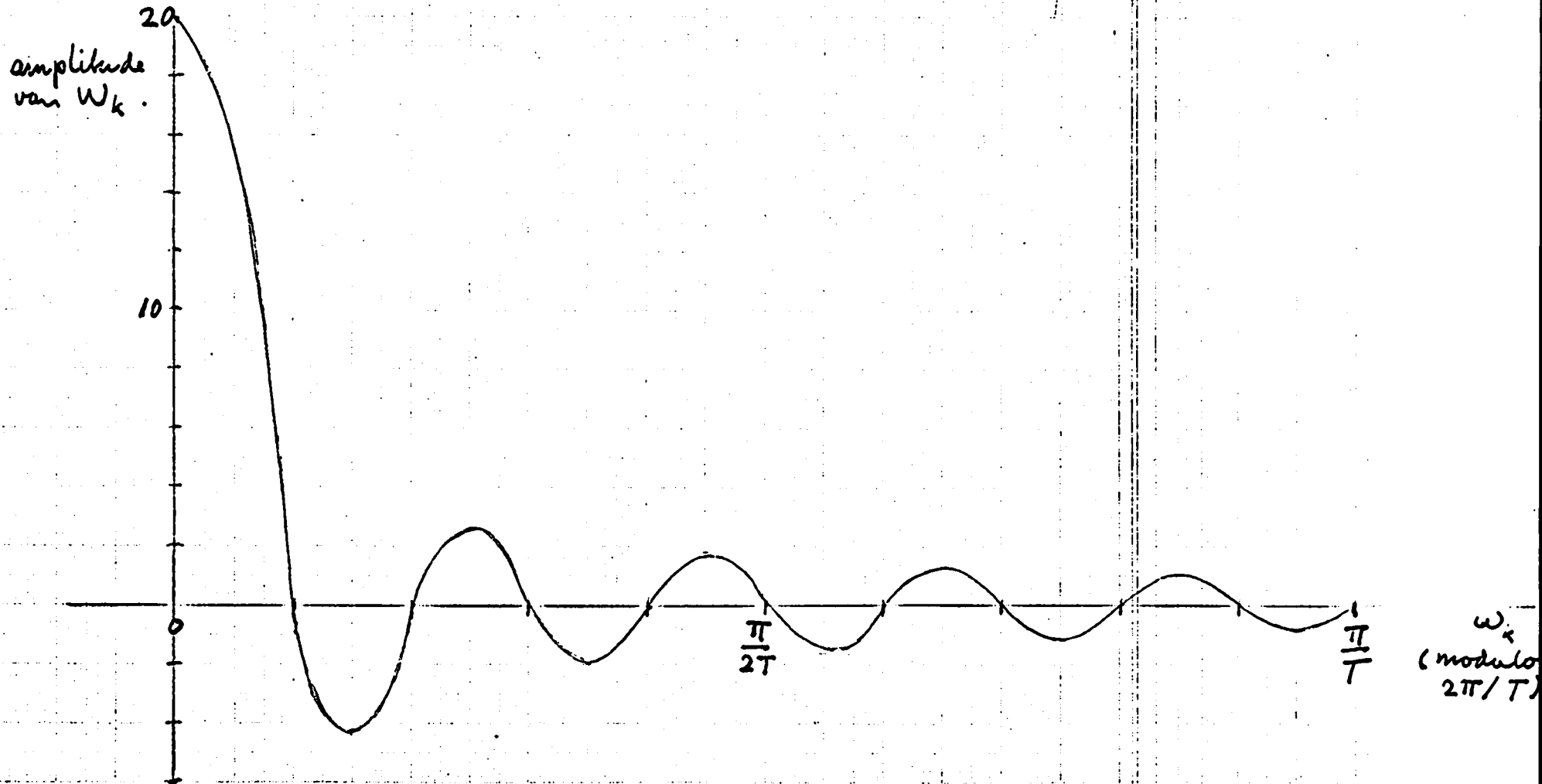
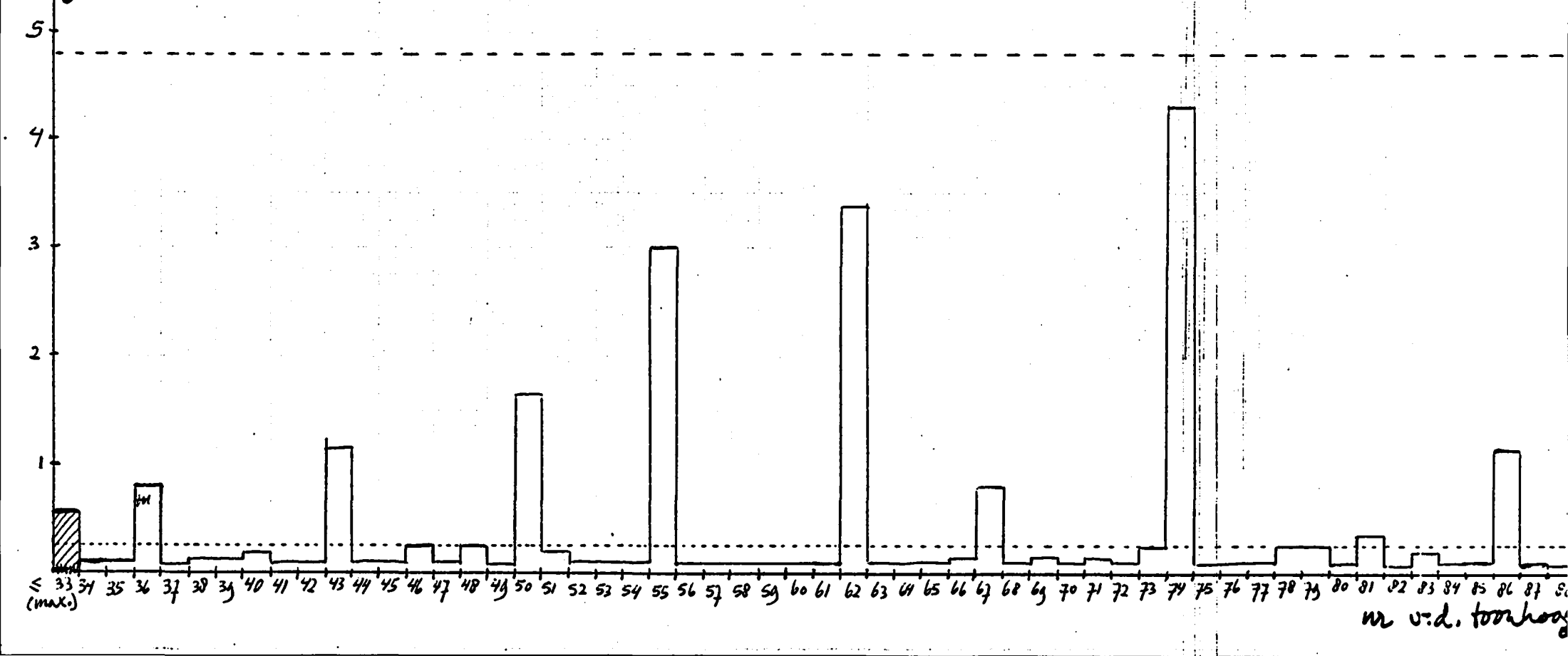


Fig. 4. De energie van $X_T(t)$ voor $1/T$ overeenkomend met de grondfrequentie van verschillende toonhoogten. (Het signaal was een zagtand spanning met toonhoogte nr 74.) Zie verder 3.5, p. 8.

Energie van $X_T(t)$.



4. Syntesemethoden.

In 2.1 gaven we als algemene beschrijving van de pianotoon:

$$S(t) = \sum_{k=1}^K a_k(t) \exp i(\omega_k t + \varphi_k).$$

Voor toepassing bij de syntese kunnen we deze beschrijving enigszins vereenvoudigen: aspecten die voor het menselijk oor niet waarneembaar zijn kunnen nu buiten beschouwing blijven. Op grond van het gestelde onder 2.4 en 2.6 kunnen we dan het reële deel van het signaal schrijven als:

$$S(t) = C(t) \sum_{k=1}^K a_k \sin \omega_k t$$

Wanneer we niet alle frequenties afzonderlijk willen schatten, kunnen we ω_k hierin bepalen uit de grondfrequentie ω_1 en de konstante B volgens

$$\omega_k = k \cdot \omega_1 \sqrt{1 + B k^2}$$

(zie 2.3).

De meest voor de hand liggende methode om een syntese van de pianosignalen te konstrueren, is nu om, wanneer een signaal eenmaal gedetekteerd is, de parameters te schatten die nodig zijn voor een beschrijving op deze manier, en vervolgens op grond daarvan het signaal samente stellen.

Wij hebben een programma geschreven wat signalen samenstelde op deze manier, zonder dat er tonen gedetekteerd of parameters geschat behoeften te worden: de beginmomenten en alle parameters werden ingelezen van ponskaarten.

Uit de resultaten van dit programma kon gekonkludeerd worden, dat de syntese van een stuk van 1 minuut op deze manier enkele etmalen rekentijd zou vergen. Een drastische bekorting van de berekeningsprocedure was dus absoluut noodzakelijk. Een stap in die richting bestaat erin, de berekeningstijd van de opeenvolgende waarden van de sinusfuncties sterk te bekorten, door hiervoor iteratieve betrekkingen te gebruiken, of door een aantal samplewaarden van 1 periode van een sinus als "stored function" in het geheugen te hebben staan.

(lit.7). Met laatstgenoemde techniek kan men nog meer tijd besparen wanneer men een signaal synteseert dat bestaat uit de (eventueel nog met een tijdsafhankelijke amplitudefaktor vermenigvuldigde) periodieke herhaling van 1 bepaalde golfvorm. Men kan dan eerst deze golfvorm berekenen en die als "stored function" in het geheugen zetten. (lit.8). Voor een signaal wat bestaat uit K componenten, wordt de rekentijd dan nog eens bijna K maal zo klein.

Voor het pianosignaal kan deze methode gebruikt worden wanneer men de inharmonisiteit van de boventonen verwaarloost. Hoewel te verwachten valt, dat dit de natuurgetrouwheid van de nabootsing enigszins schade zal doen, meenden wij toch deze concessie te moeten doen, teneinde de technische realiseerbaarheid van het geheel zo groot mogelijk te houden.

5. Schattingen, benodigd voor de syntese.

5.1. Schatting van de grondfrequentie.

De grondfrequentie van een gedetekteerde toon wordt geschat door in de buurt van de a priori geschatte grondfrequentie (met behulp waarvan de detektie tot stand gekomen is) die frequentie te zoeken waarvoor de energie van $x_T(t)$ (zie 3.5), met $T = 1/\text{frequentie}$, het grootst is.

5.2. Schatting van de golfvorm.

Als schatting van de vorm van 1 periode van het signaal gebruiken we de functie $x_p(t)$, met $T = 1/\text{grondfrequentie}$, bepaald op het tijdstip 0,07 sek. na het begin van de toon. Op dit moment is de toon nog zo luid mogelijk (d.w.z. dat de signaal-ruis-verhouding zo gunstig mogelijk is), terwijl wel de inslinger-verschijnselen afgelopen zijn. (zie 2.4)

Hierbij dient opgemerkt te worden dat de ruis die in deze schatting aanwezig is, in de gesyntetiseerde toon niet als ruis gehoord zal worden: hij wordt periodiek herhaald, en resulteert op die manier in fouten in de amplitudes van de hogere harmonischen, d.w.z. in een verandering van het timbre van de toon.

5.3. Schatting van het amplitudeverloop.

De amplitude van de toon wordt om de 512 samples (= ca 1/60 sek.) geschat uit de energie van $x_T(t)$. Tussen deze punten wordt het verloop benaderd door een lineair verloop.

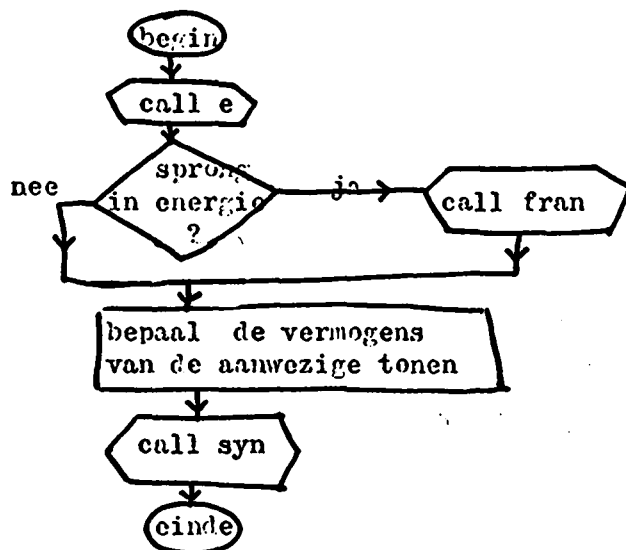
De detektie van het einde van de toon is ondergebracht bij deze schatting: het einde van de toon wordt gedetekteerd wanneer de energie van $x_T(t)$ beneden een bepaald minimum daalt.

6. Het programma filter2.

De belangrijkste aspecten van de werking van het programma zijn reeds aangegeven in par. 3.5, en in de hoofdstukken 4 en 5.

Allereerst dien hieraan te worden toegevoegd, dat we het de moeite waard hebben gevonden, de rekentijd nog aanzienlijk verder terug te brengen, door bij de detektie nog een andere bron van informatie in te schakelen dan het signaal zelf: de partituur van de gespeelde muziek. Wanneer we ervan uitgaan dat de pianist zich daaraan houdt, behoeven we alleen die tonen te zoeken die volgens de partituur verwacht kunnen worden, en kunnen we pogingen tot detektie van andere tonen achterwege laten. Als extra input van het programma gebruiken we daarom een serie ponskaarten, waarop de nummers van de achtereenvolgens te verwachten tonen zijn aangegeven.

Voor de details verwijzend naar de PI/I-tekst (Appendix I), zullen we hier volstaan met een zeer globale schets van de structuur van het programma. De achtereenvolgens te verwerken samplewaarden staan in records van 512 stuks op tape. Deze records worden 1 voor 1 verwerkt. Het flowdiagram voor de verwerking van 1 record kan als volgt worden weergegeven:



- De procedure e berekent de totale signaalenergie van de record.
- De procedure fran leest, indien nodig, het volgende stuk van de partituur in. Vervolgens wordt voor alle frequenties die volgens de partituur te verwachten zijn, nagegaan of ze 0,07 sekense deze record aanwezig zijn.
- Dit nagaan gebeurt middels de procedure check, die tevens bij een positief resultaat ervoor zorgt dat de juiste frequentie genoteerd wordt in het array van aanwezige grondfrequenties, en dat de golfvorm van het signaal op de disk geschreven wordt.

Check beslist dat een toon aanwezig is, wanneer voor $T = 1/(\text{geschatte grondfrequentie})$ de volgende criteria alle vervuld zijn:

1. Energie van x_T ligt boven een bepaald niveau. (voor betekenis x_T : zie 3.5)
2. Energie van $x_{T/2}$ ligt boven datzelfde niveau.
3. Energie van $x_{T/2}$ is kleiner dan die van x_T .

Het tweede criterium is er bij gevoerd om sommige "valse" detekties (zie b.v. in fig. 4: toon nr 67) te voorkomen.

Het derde criterium is er bij gevoerd om het detecteren van onderharmonischen van een toon (zie in fig. 4: de tonen 62, 55, 50, 43: 2e, 3e, 4e en 6e onderharmonische van de werkelijk aanwezige toon 74) te voorkomen.

- De procedure syn konstrueert voor de duur van de onderhavige record de syntese van het signaal, en schrijft het resultaat op tape2.
- De procedure son, die door verschillende van de andere procedures wordt aangeroepen, berekent 1 periode van de functie $x_T(t)$ voor $T = 1/f$.

7. De analoog-digitaal -conversie.

Voor de vervaardiging van de digitale inputband, uitgaande van een analoge bandopname van het te verwerken signaal, konden we gebruik maken van interface en software van het hybride systeem IEM 360/Hitachi.

Wanneer het signaal van de bandrecorder wordt aangesloten op een kanaal van de interface, kan men door in het programma wat op de digitale computer gedraaid wordt de procedure INREP aan te roepen, een willekeurig aantal samplewaarden in het geheugen van de digitale computer lezen. De sampletijdstippen kunnen bepaald worden door een interne of door een externe timer. Het aantal samples wat zo kan worden ingelezen wordt slechts beperkt door de geheugenruimte van de IEM 360, wat in ons geval (30000 samples/sek) een ernstige beperking is. We zullen dus steeds eerst een serie samples moeten inlezen, deze dan naar de digitale band schrijven, dan de volgende serie inlezen, enz. Om de band tijdens het wegschrijven gewoon te kunnen laten doordraaien, hebben we op een zodanige wijze van de mogelijkheid tot externe timing gebruik gemaakt, dat het hele signaal kan worden ingelezen door de band 2 maal te draaien.

De pulsen die de samplemomenten definiëren hebben we op een ander spoor gezet van dezelfde band waar het signaal op staat, en op een derde spoor een blokgolf, met een halve periode ter lengte van 512 samplemomenten. De pulsreeks en de blokgolf worden verwerkt door een elektronische schakeling die de pulsreeks alleen doorlaat wanneer de blok hoog is (eerste keer), dan wel wanneer deze laag is (tweede keer).

De samples worden zo ingelezen in records van 512 stuks. De eerste maal wordt eerst record nr 1 ingelezen en weggeschreven naar digitale tape nr 1, vervolgens wordt record nr 3 ingelezen en weggeschreven naar tape 1, dan record nr 5, enz. De tweede maal dat de band gedraaid wordt, wordt eerst record nr 1 van tape 1 naar tape 2 geschreven, dan record 2 ingelezen en naar tape 2 geschreven, dan record 3 van tape 1 naar tape 2 geschreven, enz. Aan het eind van deze procedure staan dus de records 1,2,3,4, enz. in volgorde op tape 2.

De hardware voor deze inleesmethode werd gebouwd door de heer van Nunen. De programma's vindt men in Appendix II.

Bij het verwerken van de aldus verkregen data, trad aan het licht dat er speciale problemen zitten aan het gebruik van de software van het hybride systeem vanuit PL/I. Hierop wordt gedetailleerd ingegaan in hoofdstuk 9. Daar deze problemen optraden in een kritiek laat stadium van het onderzoek, hadden ze tot gevolg dat de analysemethode van het programma filter2 niet meer getest kon worden met echte pianosignalen.

8. De digitaal-analoog-conversie.

De gang van zaken om te komen tot de convertering van de digitale outputband van het filterprogramma, stelden wij ons voor als zeer soortgelijk aan hetgeen onder 7. beschreven werd, echter in tegengestelde richting. Daar, in tegenstelling tot de verwachting, de procedure OUTREP die daarbij nodig is bij het beëindigen van ons onderzoek nog niet beschikbaar was, is dit nog niet gerealiseerd, hetgeen betekent dat we nog niet de mogelijkheid hadden, output van het filterprogramma in geluid om te zetten.

9. Het gebruik van de software van het hybride systeem in PL/I-programma's.

9.1. Inleiding.

Wanneer men in een programma dat op de digitale machine wordt uitgevoerd,, informatietransport van en/of naar de analoge computer (of b.v. een analoge bandrecorder) wenst te incorporeren, kan men daartoe gebruik maken van een voor dit doel geschreven software-pakket. Het betreft hier subroutines, geschreven in assembler-taal, en bedoeld om aangeroepen te worden in FORTRAN-programma's. Deze subroutines kan men ook in PL/I-programma's aanroepen, doch men moet dan bedacht zijn op de consequenties van het volgende: I.h.a. worden de input-data door de subroutine geacht als halfwoorden in het geheugen te staan. De output-data worden als halfwoorden in het geheugen gezet. Deze halfwoorden bestaan uit 2 tekenbits, gevolgd door 14 bits, die de absolute waarde van het getal binair representeren. Arithmetische data worden in PL/I echter nooit op deze manier gecodeerd: binaire getallen worden voorgesteld door een heel woord, bestaande uit v.l.n.r. 1 tekenbit en 31 bits, die de absolute waarde van het getal binair weergeven; voor decimale getallen is de z.g. "packed decimal" codering van toepassing. (lit. 9).

9.2. Arithmetische parameters van de subroutines.

9.2.1. Algemeen.

Het voorgaande heeft tot gevolg dat men bij de specificatie van arithmetische parameters niet zonder meer de gewenste waarde op kan geven, doch ervoor zorg moet dragen dat deze waarde op de juiste wijze gecodeerd staat in de eerste twee bytes van de opgegeven variabele of konstante. Bij positieve getallen kan men dit doen door een fixed binary getal te specificeren dat een 2^{16} maal te grote waarde heeft. Bij negatieve getallen moet men technieken gaan toepassen zoals we onder 9.3 aangeven.

9.2.2. "Blauwe" subroutines.

Van de procedure DUMP bestaat een versie op blauwe kaarten, die de beide arithmetische parameters (die resp. het aantal cycli en het nr van het kanaal bepalen) leest in de tweede helft van het opgegeven (hele) woord. Omdat beide getallen altijd positief zijn, is het dan voldoende om te zorgen dat ze 'fixed binary' gecodeerd zijn.

9.3. ~~Output- en input-arrays van de subroutines.~~

Wanneer de subroutine als output een array met getallen geeft, dan staan deze in het geheugen in opeenvolgende halfwoorden. Halfwoorden kunnen in PL/I slechts bit-strings zijn. B.v. "bit(16)". Wanneer men deze bit-strings in arithmetische expressies laat voorkomen, worden ze gekonverteerd tot binaire getallen. Bij de konversie wordt de bit-string geïnterpreteerd als de binaire representatie van een positief getal, zonder tekenbit.

Tekenbits op de eerste plaatsen van de eerste byte leiden dus tot fouten: wanneer er negatieve getallen voorkomen kan men niet rekenen met bit-strings. In dat geval moet men de halfwoorden converteren tot hele woorden, zodanig dat: de eerste bit van het hele woord = de eerste bit van het halve woord, de laatste 14 bits van het hele woord = de laatste 14 bits van het halve

woord, de tussenliggende bits van het hele woord zijn nul. Deze halfwoorden schrijft men dan op disk of tape, of assigneert ze (n.b.v. de pseudo-variabele UNSPEC) aan de elementen van een "fixed binary" array. De genoemde wijze van conversie kan men in PL/I realiseren n.b.v. de functie en pseudo-variabele SUBSTR. Ons is gebleken dat de conversie van 1 halfwoord op die manier 12 msec duurt. De conversie van een signaal van 1 minuut, gesampled met 30 kHz, kost dan 6 uur, wat onpraktisch lang is. Wanneer er hiervoor in PL/I niet iets snellers gevonden kan worden, zullen er aparte subroutines voor gebruik in PL/I nodig zijn.

10. Konklusies.

Op verschillende punten dient het onderzoek voortgezet te worden:

- Allereerst dient de input- en output- software volledig bruikbaar gemaakt te worden.
- Dan dient het filterprogramma aan de praktijk getoetst te worden, waarbij zal blijken op welke punten onze oplossing niet bevredigend is, en gecorrigeerd of geheel verworpen dient te worden.
- Ook zou een zuiver theoretisch onderzoek over de optimale methode om te discrimineren tussen signalen van het door ons beschouwde type, een betere basis aan het verdere werk kunnen leveren.

11. Literatuur.

1. O.H. Schuck en R.W. Young: Observations on the vibrations of piano strings.
J.A.S.A. 15, 1 (juli '43), p. 1.
2. H. Meinel: Musikinstrumentenstimmungen und Tonsysteme.
Acustica 7 ('51), p. 185.
3. J. Meyer en W. Lottermoser: Ueber die Moeglichkeit einer klanglichen
Beurteilung von Fluegeln.
Acustica 11 ('61), p. 291.
4. H. Fletcher, E.D. Blackham en R. Stratton: Quality of piano tones.
J.A.S.A. 34, 6 (juni '62), p.749.
5. P. Nykhoff: Informatietheorie. Collegesaantekeningen '66/'67. Groep M
Meten & Regelen, afd. E, T.H.E.
6. Collected algorithms from CCM, nr 338 en 339.
7. Collected algorithms from CCM, nr 345.
8. M.V. Mathews: The Technology of Computer Music.
1969, MIT-Press, Cambridge (Mass.) & London.
9. IBM System/360. PL/I subset reference manual. (IBM systems reference library.)