

**KONINKLIJK NEDERLANDS
METEOROLOGISCH INSTITUUT**

VERSLAGEN

V-335

W. Kuik *

**Inleiding in enkele methoden voor
het schatten van spectra.**

- * Dit verslag resulteerde uit een literatuurstudie op het gebied van het bepalen van energiespectra van eindige tijdreeksen en werd gedaan op het K. N. M. I. door Wouter Kuik, in een stageperiode van drie maanden in 1979 (1 februari - 1 mei), o.l.v. Drs. A. G. M. Driedonks.

De Bilt 1980

Publikatienummer: K.N.M.I. V-335 (MO)

Koninklijk Nederlands Meteorologisch Instituut,
Meteorologisch Onderzoek,
Postbus 201,
3730 AE De Bilt,
Nederland.

U.D.C.: 518.5

INHOUDSOPGAVE

	<u>blz.</u>
Lijst van symbolen, afkortingen en tekens	1
Inleiding	2
Kort historisch overzicht	3
Deel I: Belangrijke definities en eigenschappen van spectra	6
1. Definitie van het energiespectrum	6
2. Veronderstellingen over de autocovariantie functie	9
3. Het probleem van de stationariteit van het stochastisch proces	12
4. De eindige lengte van een tijdreeks	13
5. Veronderstelling over de gegevens buiten het meetgebied	15
6. Vensters en filters	16
7. Aliasing	21
8. Referenties bij inleiding en deel I	22
Deel II: Bespreking en vergelijking van verschillende methoden voor het berekenen van spectra	23
9. De periodogramschatter	23
10. De venstermethoden	26
11. De Variabele Record Length methode	30
12. De maximale entropie methode (MEM)	31
13. De exacte spectraalrepresentatie	44
14. Het bepalen van een spectrum door filteren	47
15. Vergelijking van de verschillende methoden	50
16. Tabel van vergelijkende maatstaven	51
17. Toepassingsgebieden en conclusies	52
18. Referenties bij deel II	54
A1 Aanhangsel	57

Lijst van symbolen, afkortingen en tekens.

t	: De tijdsvariabele.
τ	: De tijdverschuivingsvariabele in het convolutie- domein.
f (Hz)	: De frequentie.
ω (rad/s)	: De hoeksnelheid = $2\pi f$.
$X(t)$: Het stochastische proces.
$x(t)$: Een realisatie van $X(t)$ gemeten in eindig tijds- interval.
$E(x)$: De verwachtingswaarde van x .
$\mu = E(X(t))$: De verwachtingswaarde van $X(t)$.
$F_o(x)$: De Fourier getransformeerde van x .
$*$: Het convolutieteken, zodat $(x * y)(\tau) =$ $E(x(t).y(t-\tau))$.
$\gamma(\tau)$: De autocovariantiefunctie (a.c.v.f.) van $X(t)$ is $E((x(t)-\mu) \cdot (x(t-\tau)-\mu))$.
$\rho(\tau)$: De autocorrelatiefunctie (a.c.f.) van $X(t)$ is $\gamma(\tau)/\gamma(0)$.
$c(\tau)$: De a.c.v.f. van $x(t)$.
$\Gamma(f)$: Het zuivere energiespectrum van $X(t)$ is $F_o(\gamma(\tau))$.
$C(f)$: Het onzuivere spectrum van $X(t)$ is $F_o(c(\tau))$.
$c_a(\tau)$: De zuivere maar afgebroken schatter voor $\gamma(\tau)$.
T	: De eindige meettijd van monster $x(t)$ van $X(t)$.
N	: Het aantal samples (monsters) in de tijd T .
Δ	: Het bemonsteringsinterval.
f_N	: De Nyquistfrequentie is $1/(2\Delta)$.
F.F.T.	: Fast Fourier Transformatie.
D.F.T.	: Discrete Fourier Transformatie.

- $w(\tau)$ of $w(t)$: Het venster in het convolutie- of tijdsdomein.
- $W(f)$: Het spectrale venster in de genormeerde $F_0(w(\tau))$.
- F^* : De complextoegevoegde van F_0 zodat $F.F^* = [F]^2$
- $\bar{c}(\tau)$: De a.c.v.f. $\gamma(\tau)$ gezien door venster met lengte $M < T$.
- $\bar{C}(f)$: Het onzuivere, gesmoothte spectrum $F_0(\bar{c}(\tau))$.
- M : De afbreeklengthe of maximale tijdverschuiving in τ -domein.

Inleiding.

Een literatuurstudie werd gedaan op het terrein van het bepalen van spectra van eindige tijdreeksen.

Het verslag wil de lezer helpen bij het bepalen van een energiespectrum uit een tijdreeks. Tot zijn beschikking staan verscheidene methoden, ieder met eigen voor- en nadelen.

De klassieke periodogram- en filter methoden (Blackman en Tukey, 1958) zijn het meest bekend, maar vertonen soms grote onvolkomenheden, vooral in extreme omstandigheden, zoals pieken in het spectrum en bij korte tijdreeksen. Modernere methoden, zoals de autoregressive, maximale entropie methode (M.E.M.), worden beschreven. Deze hebben een zeer hoog oplossend vermogen. Ook wordt het digitaal filteren behandeld.

Het verslag is gesplitst in twee delen. Deel I begint na een kort historisch overzicht en behandelt de problematiek, die algemeen bij het bepalen van een spectrum aan de orde komt.

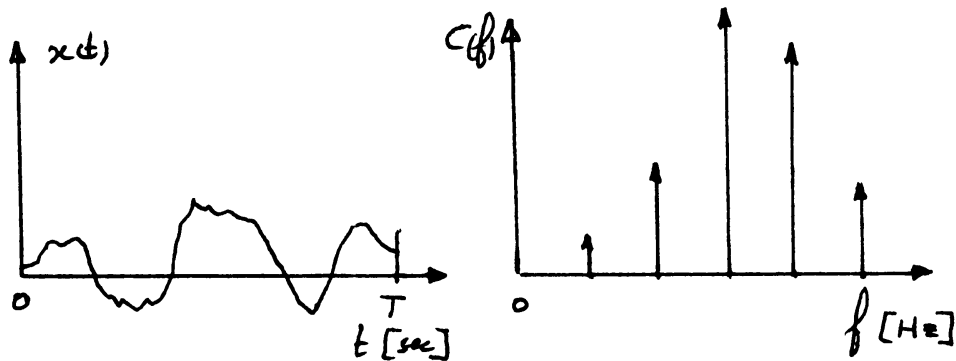
Deel II behandelt de specifieke methoden met hun voor- en nadelen.

Er wordt getracht het principe van de methoden te belichten en een inleiding te geven in de besproken methoden. Referenties geven de weg aan naar verdere kennis. Bekeken worden: de periodogram methoden met of zonder venster; de variabele record length (VRL) methode; de maximale entropie methode (M.E.M.); de maximum likelihood methode (M.L.M.); de exacte spectraal representatie en methoden gebaseerd op filteren, zowel analoog als digitaal.

Aan het eind van deel II worden de methoden vergeleken. Een aanbeveling wordt gedaan t.a.v. het toepassen van de meest geschikte methode voor het bepalen van het energiespectrum uit een gegeven eindige tijdreeks.

Een kort historisch overzicht.

De weergave van een eindige tijdreeks $x(t)$, $t \in (0, T)$ in het spectrale domein gebeurde eerst door middel van een eenvoudige discrete ontbinding in sinussen en cosinussen. De energie, die elke frequentiecomponent bevatte werd weergegeven in een stakendiagram, het zogenaamde periodogram (Schuster, 1898).



Figuur 1. De tijdreeks $x(t)$ met zijn periodogram $C(f)$, berekend volgens:

$$C_k = 2 \left[\left\{ \frac{1}{2n} \sum_{m=1}^{2n} x_m \cos \frac{2\pi km}{2n} \right\}^2 + \left\{ \frac{1}{2n} \sum_{m=1}^{2n} x_m \sin \frac{2\pi km}{2n} \right\}^2 \right].$$

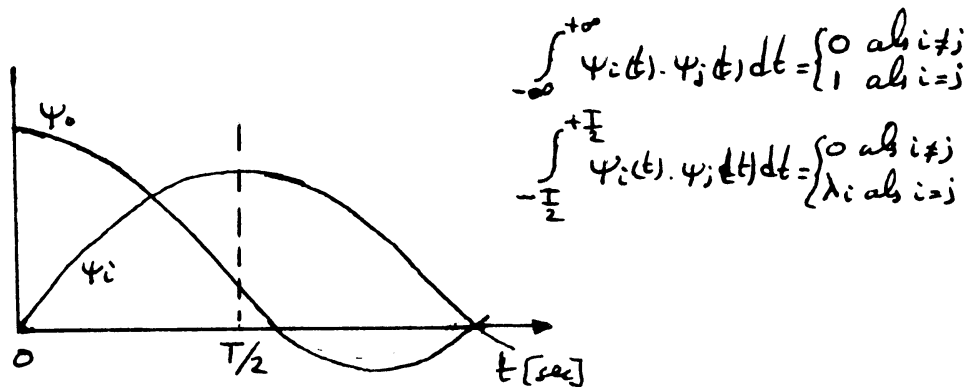
met $x(t)$ bemonsterd tot $2n$ waarden x_m en $k = 0, 1, \dots, n$.

Elke schatter $C(f)$ in dit periodogram is behept met een grote statistische fout. Door het ruwe, piekige periodogram te bewerken met een "smoothing" venster in het spectrale domein, wordt een beter toonbaar, gladder, spectrum gevonden, dat kleinere variantie heeft als $C(f)$, maar dat bij pieken een fout vertoont.

Dit werd het eerst gedaan door Hamming in 1948 (Blackman en Tukey, 1958). Algemeen wordt het toepassen van een venster de Blackman en Tukey methode genoemd. Met de opkomst van de digitale computer werd het digitale filteren ontwikkeld als hulptechniek door o.a. Radar en Gold (1965).

Snelle algoritmes zoals de Fast Fourier Transform (F.F.T.) werden toen ook ontwikkeld (Cooley, Lewis, Welch, 1967).

Wat later werden methoden gebaseerd op het optimaal gebruik van de aanwezige informatie ontwikkeld, de autoregressive methoden (A.R.), waaronder de maximum likelihood (M.L.M.) en de maximale entropie methode (M.E.M.). Pioniers op dit gebied waren V.d. Bos (1970), Burg (1967) en Ulrich (1972). De M.E.M. heeft een hoog oplossend vermogen en is tegenwoordig als computerroutine bij vele instituten aanwezig. Vergeten is een techniek, die de exacte spectraal representatie genoemd wordt (Zagalsky, 1967). Deze is gebaseerd op het ontbinden van een eindige stationaire tijdreeks in spheroïde golf functies, die de eigenschap hebben, dat ze orthogonaal zijn, zowel in één eindige tijd, als over oneindige tijd. Deze methode wordt in dit rapport weer naar voren gehaald.



Figuur 2. De spheroïde golf functies ψ_i met dubbele orthogonaliteit.

De periodogram- en filtermethode geven de lage frequenties slecht weer, vooral omdat weinig discrete lage frequenties worden meegenomen. Om hieraan tegemoet te komen werd door Schikedanz en Bowen (1975) de variabele record length techniek ontwikkeld, die bestaat uit het herhaald toepassen van een discrete Fourier routine over een verkleinde en vergrote tijdreekslengte. Dit kost echter erg veel rekentijd. Jenkinson (1977) past een logaritmische verdeling van de frequenties toe. De grote golfgetallen komen dan beter tot hun recht.

Deel I: Belangrijke definities en eigenschappen van spectra.

1. De definitie van het energiespectrum.

Belangrijk is, hoe het spectrum wordt gedefinieerd. In de literatuur is soms verwarring te bespeuren. Dit komt door het feit, dat het spectrum, bepaald uit de eindige reeks, verschillend is van het echte spectrum, dat immers voor de oneindige reeks is gedefinieerd.

Het "echte" spectrum $\Gamma(f)$ geeft de verdeling weer van de scalaire grootheid energie in het spectrale domein, dus hoe deze over de frequentie f is verdeeld.

Met "de energie" wordt bedoeld het ensemble gemiddelde van energie $E([x(t)]^2)$ over de realisaties van lengte T . Bij een stationair stochastisch proces $X(t)$ is dit ensemble gemiddeld gelijk aan het tijdsgemiddelde over oneindige tijd.

Dit betekent, dat in principe slechts van een stationair stochastisch proces het spectrum kan worden bepaald, dat bovendien karakteristiek is voor dit proces (of het instationaire proces moet herhaalbaar zijn). Als $C(f)$ de verdeling is van energie van $x(t)$ met $t \in (-T/2, T/2)$, dan kan het spectrum geschreven worden (Jenkins en Watts, 1962).

$$\Gamma(f) = \lim_{T \rightarrow \infty} E(C(f)).$$

Zowel het spectrum $\Gamma(f)$ als het "sample" spectrum $C(f)$ kunnen gezien worden als Fourier getransformeerden van respectievelijk de autocovariantiefunctie $\gamma(\tau)$ en de "sample" autocovariantiefunctie $c(\tau)$.

$$\Gamma(f) = \int_{-\infty}^{+\infty} \gamma(\tau) e^{-j2\pi f \tau} d\tau,$$

$$\text{en } C(f) = \frac{1}{T} \left[\int_{-\frac{T}{2}}^{\frac{T}{2}} x(t) \cdot e^{-2\pi j f t} dt \right]^2 =$$

$$\int_{-T}^{+T} c(\tau) \cdot e^{-2\pi j f \tau} d\tau. \quad (1.1)$$

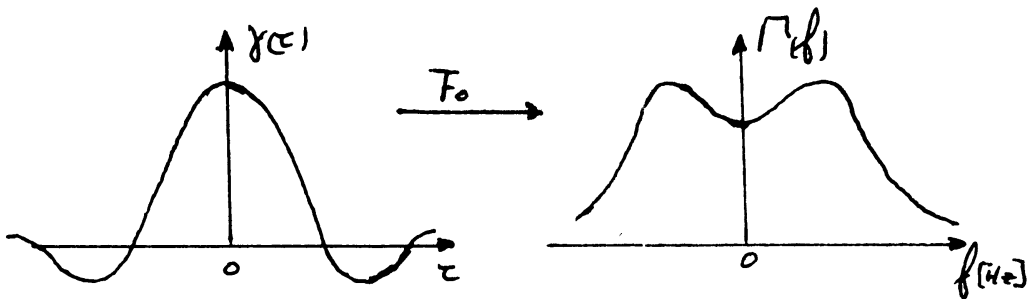
$$\text{met } c(\tau) = \frac{1}{T} \int_0^{T-|\tau|} (x(t) - \bar{x}) \cdot (x(t + |\tau|) - \bar{x}) dt$$

$$= 0$$

voor $0 < |\tau| < T$
voor $|\tau| > T$

Deze sample a.c.v.f. is dus nul als $|\tau| = T$. (Jenkins en Watts, 1968).

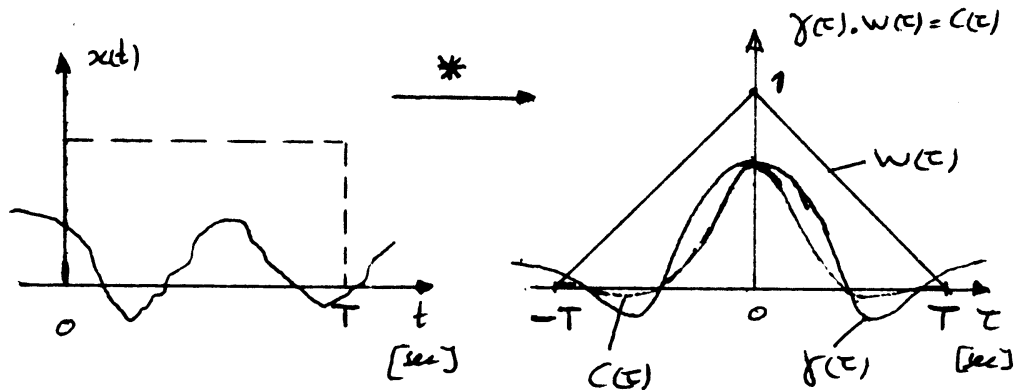
Het werkelijke spectrum $\Gamma(f)$ wordt bepaald door Fouriertransformatie van de symmetrische oneindige a.c.v.f. $\gamma(\tau)$.



Figuur 3. $\gamma(\tau)$ en zijn Fouriergetransformeerde, het spectrum $\Gamma(f)$.

Als over het interval $t = \{0 \dots T\}$, $x(t)$ gemeten wordt, betekent dat een rechthoekig tijdvenster over $x(t)$ en dus in het convolutiedomein (τ) een driehoekig (Bartlett) venster $w(\tau)$ over de a.c.v.f. $\gamma(\tau)$.

De vermenigvuldiging van $\gamma(\tau)$ met dit venster wordt gemeten en heet de sample a.c.v.f. $c(\tau)$, die slechts bekend is over $|\tau| < T$.

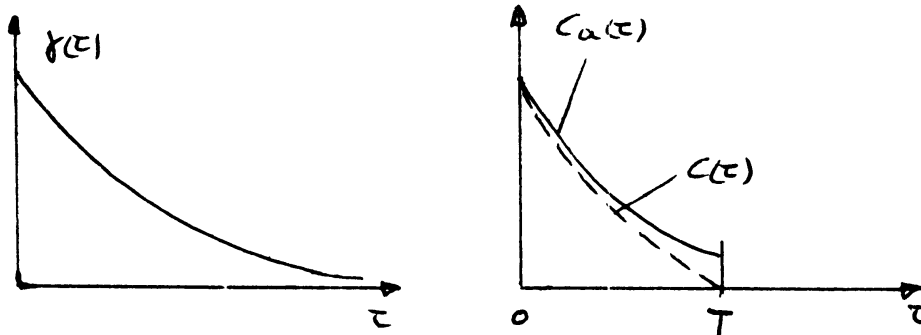


Figuur 4. Als $X(t)$ over eindige tijd bekeken wordt, betekent dit dat de a.c.v.f. $\gamma(\tau)$ door een driehoekig (Bartlett) venster $w(\tau)$ gezien wordt als de vervormde, onzuivere, afgebroken $c(\tau) = w(\tau) \cdot \gamma(\tau)$.

Als T nadert tot oneindig zien we, dat dit driehoekig venster, een rechthoekig oneindig venster wordt, zodat $c(\tau)$ gelijk $\gamma(\tau)$ is. $c(\tau)$ heet een asymptotisch zuivere schatter van $\gamma(\tau)$. Echter bij beperkte tijd T is $c(\tau)$ een onzuivere (biased) schatter. Een zuivere, maar afgebroken schatter voor $\gamma(\tau)$ kan eenvoudig worden gevonden door $c(\tau)$ door het venster $w(\tau)$ te delen; dit komt neer op het bepalen van $c_a(\tau)$ (Jenkins en Watts, blz. 174-175).

$$\text{met } c_a(\tau) = \frac{c(\tau)}{1 - \frac{|\tau|}{T}} = \frac{1}{T - |\tau|} \int_0^{T-|\tau|} (x(t) - \mu)(x(t + |\tau|) - \mu) dt$$

voor $0 < |\tau| < T$



Figuur 5. De a.c.v.f. van een stochastisch proces $S(t)$ en de sample a.c.v.f. $c_a(\tau)$ van realisatie $x(t)$ met $\mu = 0$.

Noot. De variantie van $c_a(\tau)$ wordt steeds groter naarmate $\tau \rightarrow T$, omdat steeds minder termen worden meegenomen. Bij $c(\tau)$ wordt dit gecompenseerd door vermenigvuldiging met $w(\tau)$.

De $c(\tau)$ of zijn getransformeerde, het samplespectrum $C(f)$ zijn discreet te bepalen uit de tijdreeks $(0 \dots T)$ en vormen de basis voor de spectrale methoden beschreven in deel II. $C(f)$ is de ruwe periodogram-schatting van $\Gamma(f)$.

Bij verscheidene methoden worden veronderstellingen gedaan t.a.v. het verloop van de a.c.v.f. $c_a(\tau)$. Deze veronderstellingen worden nu kort behandeld.

2. Veronderstellingen over de autocovariantiefunctie.

Geen veronderstelling wordt gedaan bij de ruwe periodogram-schatting. De schatting is gelijk aan de Fouriergetransformeerde van de biased a.c.v.f. $C(\tau)$ en kan direct of indirect uit de tijdreeks $(0 \dots T)$ berekend worden volgens (1.1).

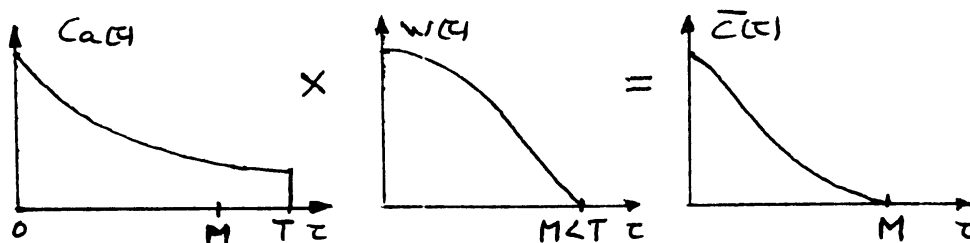
Directe D.F.T. $C(f) = \frac{1}{T} \left[\int_{-\frac{T}{2}}^{+\frac{T}{2}} x(t) \cdot e^{-2\pi jft} dt \right]^2 =$

$$\frac{1}{T} Fo^2 \{x(t)\}. \tag{2.1}$$

Indirect $C(f) = 2 \int_0^T c(\tau) e^{-2\pi jf\tau} d\tau = 2Fo \{c(\tau)\}. \tag{2.2}$

Beide resultaten zijn identiek; de methoden zijn gelijkwaardig. Terugtransformatie van $C(f)$ levert de $c(\tau)$ weer op, echter niet de oorspronkelijke tijdreeks $x(t)$, daar fase-informatie door het kwadrateren verloren is gegaan.

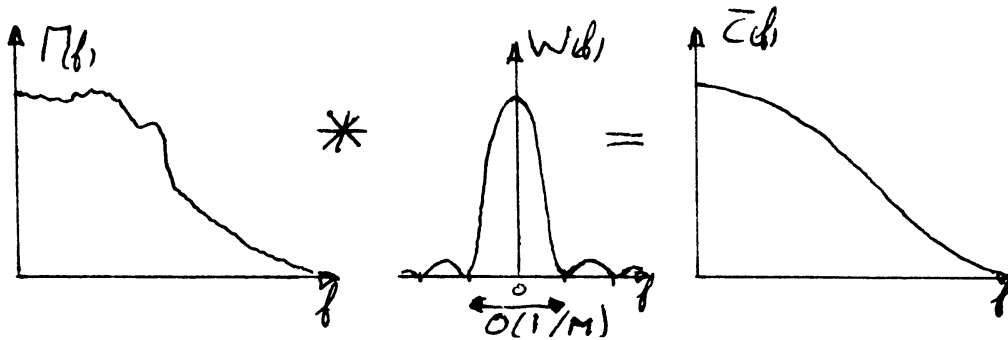
Bij de venstermethode wordt de sample a.c.v.f. $C_a(\tau)$ naar nul gebracht door een vermenigvuldiging met een venster $w(\tau)$ in het convolutiedomein. De toepassing van zo'n venster kan in het tijdsdomein worden uitgelegd als het toepassen van een taper (vermenigvuldiging van $x(t)$ met taperfunctie over $t = 0 \dots T$). De discontinuïteit in $x(t)$ door het rechthoekig tijdvenster wordt hierdoor verminderd, met als gevolg minder harmonischen in het spectrum (zie deel II, het ruwe periodogram).



Figuur 6. De Blackman en Tukey-methode, vermenigvuldiging met convolutievenster $w(\tau)$.

Deze vermenigvuldiging levert een foutieve $\bar{c}(\tau)$, die nul is voor $\tau = M \leq T$ en die getransformeerd een biased, maar gladder (kleinere variantie) spectrum $\bar{C}(f)$ oplevert.

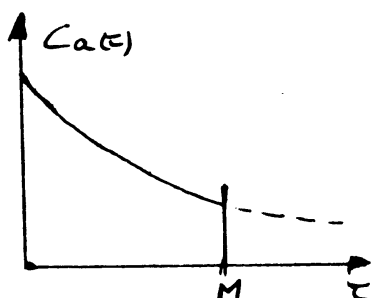
Vermenigvuldiging in het convolutiedomein met venster $w(\tau)$ van $\gamma(\tau)$ of $c_a(\tau)$ is in het spectrale domein convolutie of "smoothing" met een spectraalvenster (of filter) $W(f)$ van $\Gamma(f)$.



Figuur 7. Convolutie van spectrum Γ met filter $W(f) = F_0 \{w(\tau)\}$ levert het biased spectrum $\bar{C}(f)$ (met oneindige bandbreedte, doordat $W(f)$ oneindig is in f).

Verscheidene soorten vensters kunnen worden toegepast (zie deel II, venstermethoden), ieder met eigen fout en smoothing (bandbreedte).

Bij de autoregressieve methoden wordt de afgebroken a.c.v.f. $c_a(\tau)$ geëxtrapoleerd. Bij de Maximum entropie methode is de voortzetting er één waarvan de gemiddelde informatie per tijdseenheid maximaal is. De gemiddelde informatie per tijdseenheid wordt entropie genoemd. De M.E.M. levert inderdaad de beste resultaten. Burg (1972) toonde aan dat het spectrum Γ een maximale entropierepresentatie is van de energie verdeling van de stochastische variabele $X(t)$ (zie deel II, blz. 32).



Figuur 8. Uitgaande van een goede $c_a(\tau)$, $|\tau| < M < T$ wordt met dezelfde of kleinere variantie dan $c_a(M)$ bij de M.E.M. of M.L.M. het verloop van de a.c.v.f. gecontrueerd. F_0 transform levert dan het spectrum.

Bij de exacte spectraal representatie, wordt het onbekende deel van $\gamma(\tau)$ dat buiten het blok $|\tau| < M$ ligt, met behulp van spheroïde golf-functies exact geconstrueerd. Het deel van $\gamma(\tau)$ binnen het blok, wordt daartoe eerst ontbonden in deze golf-functies en extrapolatie levert dan het onbekende deel van $\gamma(\tau)$.

Er worden hier dus geen veronderstellingen gedaan, die fouten veroorzaken in het spectrum Γ .

3. Het probleem van de stationariteit van het stochastische proces.

Bij alle methoden ter bepaling van een schatter $C(f)$ van het spectrum $\Gamma(f)$ wordt stationariteit van het stochastische proces $X(t)$ verondersteld. Dit betekent dat $C(f)$ gemeten over eindige tijd T representatief is voor andere tijdsintervallen en ook voor het stochastische proces.

Bij een stationaire reeks is de verdeling van energie in het spectrale domein, over grote tijd gemeten, voor iedere tijdsinterval T gelijk. Dit betekent niet, dat van een niet stationair proces $X(t)$ geen spectrum kan worden bepaald. Echter dit spectrum is slechts een representatie van het tijdsinterval waarover gemeten is; d.w.z. de voorspellende eigenschappen zijn beperkt.

Een goede test voor stationariteit is vaak moeilijk te geven, omdat in de praktijk slechts de reeks over eindige tijd bekend is, zoals b.v. bij meteorologische reeksen, waar de lage golfgetallen vaak sterk bijdragen in de energie (klimaatschommelingen).

Wat betreft stationariteit lijken er twee mogelijkheden:

- a. Het is bekend dat het proces stationair is vanuit fysische redenen; gezocht wordt nu een tijdsinterval, dat groot genoeg is om representatief te zijn met zekere nauwkeurigheid (variatie) voor het oneindige stationaire proces. Dit interval is afhankelijk van de gevolgde methode en natuurlijk van de verdeling van de energie van de lage frequenties. Om een stationair proces met zekere nauwkeurigheid met een bepaalde methode te beschrijven is een minimale (methode afhankelijk) meettijd of reekslengte vereist. Aan het eind van deel II worden hierover aanbevelingen gedaan.

- b. Het is bekend dat het proces instationair is, ook vanuit fysisch oogpunt; gezocht wordt nu een tijdsinterval dat klein genoeg is om het instationaire proces als stationair te benaderen.
Als een spectrum van een niet stationair proces wordt bepaald, wordt het spectrum over een eindige meettijd representatief (stationair) verondersteld.

4. De eindige lengte van een tijdreeks.

Als van een stationair proces $X(t)$ een monster wordt genomen over het eindige tijdsinterval $(-\frac{T}{2}, +\frac{T}{2})$, wordt het proces $X(t)$ als het ware door een tijdsvenster bekeken. Een vermenigvuldiging treedt op van de oneindige reeks $X(t)$ met het eindige rechthoekige tijdvenster $w(t)$. In het spectrale domein betekent dit convolutie of smoothing van het spectrum $\Gamma(f)$ van de oneindige reeks $X(t)$ met het genormeerde kwadraat van de Fouriergetransformeerde, $W(f)$ van $w(t)$. Het spectrale venster is de Fouriergetransformeerde van het driehoekige convolutie venster en heeft de vorm van een Bartlett spectraal venster.

$$W_B(f) = T \left(\frac{\sin \pi fT}{\pi fT} \right)^2. \quad (\text{Jenkins en Watts, 1968})$$

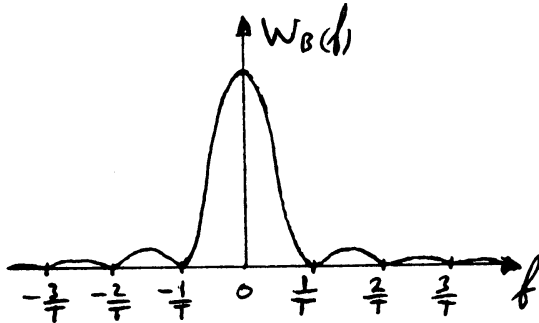
Afgevlakt wordt dus over de bandbreedte (zie verder) $B = \frac{2}{T}$ van het spectrale venster. Het oplossend vermogen behorende bij de meettijd T wordt dus bij toepassing van zo'n convolutie van Γ met W_B bepaald door deze bandbreedte.

$$C(f) = \int_{-\infty}^{+\infty} W_B(g) \cdot \Gamma(g - f) dg.$$

$C(f)$ heeft dus een onderscheidingsvermogen in frequenties van $\Delta f = \frac{2}{T}$, dus afhankelijk van de meettijd.

De functie $W(g)$ is als Fouriergetransformeerde van het eindige convolutie venster $w(\tau)$ een functie die oneindig is, d.w.z. in feite oneindige bandbreedte.

$$W_B(f) = \int_{-T}^{+T} w(\tau) e^{-2\pi j f \tau} d\tau:$$



Figuur 9. De oneindige breedte functie $W(g)$ met zijlobben.

Dit heeft tot gevolg dat $C(f)$ principieel een oneindige bandbreedte heeft, hoewel voor hoge frequenties de bijdragen minimaal worden als f toeneemt. Met toenemende meettijd T wordt de bandbreedte van C dus kleiner, binnen de meetnauwkeurigheid.

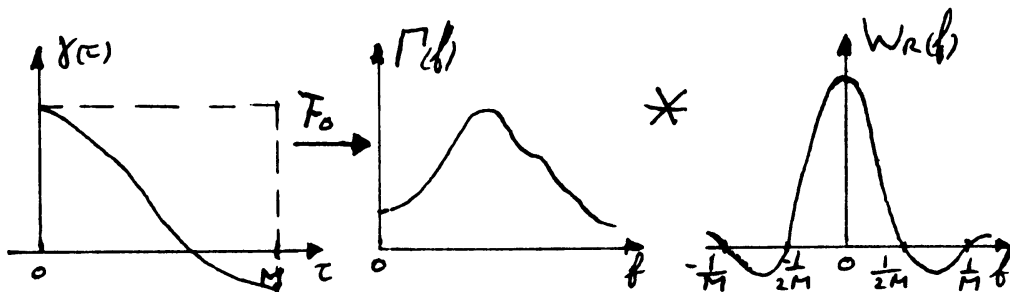
Een rechthoekig convolutievenster over $\gamma(\tau)$ betekent in het spectrale domein convolutie met $W_r(f)$:

$$W_r(f) = 2M \left(\frac{\sin^2 \pi f M}{2 \pi f M} \right)$$

(Jenkins en Watts, blz. 244).

M is hier de afbreek lengte waarbuiten de a.c.v.f. nul is.

De hoofdlob heeft breedte $1/M$.



Figuur 10. Het afbreken van de a.c.v.f. $\gamma(\tau)$ levert getransformeerd, de geconvolveerde van $\Gamma(f)$ met $W_r(f)$, met een breedte van de hoofdlob van $\frac{1}{M}$.

$$C(f) = \int_{-\infty}^{+\infty} W(g - f) \Gamma(g) dg.$$

Op deconvolutie van deze vergelijking met $W(f) = \frac{\sin M 2 \pi f}{\pi f}$ is de exacte spectraal representatie methode gebaseerd.

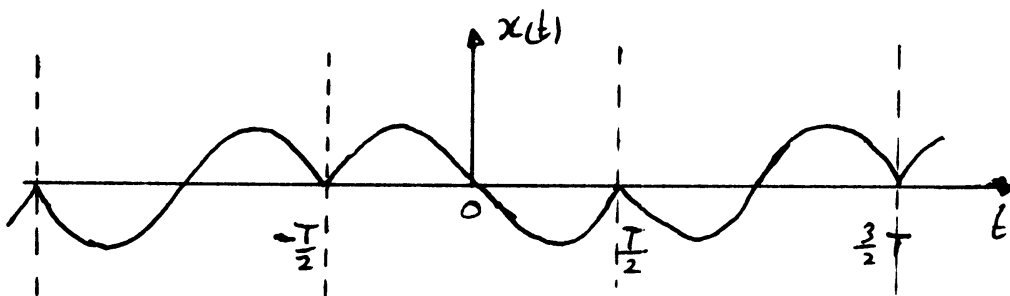
Het oplossend vermogen wordt nu niet meer beperkt door de afbreek- lengte of meettijd, maar slechts de hoeveelheid spheroïde golf functies die worden meegenomen, dus door rekentechnische aspecten.

Dit is vooral voor de lage frequenties van belang, daar de laagst te detecteren frequentie gelijk is aan de bandbreedte van het spectrale venster, bij de conventionele methoden gelijk orde ($1/T$ of $1/M$).

5. Veronderstelling over de gegevens buiten het meetgebied,

Bij de periodogram- en venstermethode wordt de reeks voor eindige Fouriertransformatie nul verondersteld, of beter gelijk aan het ge- middelde, buiten het meetgebied.

De Fouriertransformatie is eindig in de tijd en is identiek aan een oneindige Fouriertransformatie over de steeds gespiegelde eindige reeksen over $-\frac{T}{2} < t < \frac{T}{2}$. Zie voor een sinus bijvoorbeeld figuur 11.



Figuur 11. Eindige Fouriertransformatie over T is gelijk aan oneindige Fouriertransformatie over de steeds gespiegelde oneindige reeks.

Dit heeft tot gevolg, dat er in het spectrum een periodiciteit met golf- lengte $2T$ voorkomt. Tevens is door het afbreken van de reeks de band- breedte oneindig, d.w.z. door het afbreken worden hogere harmonischen geïntroduceerd.

Toman (1965) toonde door numeriek doorrekenen van een eindige Fourier- transformatie aan, dat door het afbreken ook verschuivingen van bij- dragen in het spectrale domein voorkomen, die vooral ernstig zijn bij grote golfgetallen van de orde van meettijd T .

Als na één periode een sinus wordt afgebroken, zoals in figuur 11, treedt er een verschuiving op van de piek van 16% naar lager frequenties. Golfgetallen groter dan $1,72 T$ worden als gelijkspanning gedetecteerd.

Met uitzondering van de exacte spectraal methode, zijn geen van de methoden in staat, frequenties lager dan $0,58/T$ te detecteren en dan nog met een fout.

Frequenties boven de $4/T$ worden maximaal met een verschuiving van 5% in lagere frequenties gedetecteerd.

Zie voor de M.E.M. Chen en Stegen (1974). De M.E.M. vertoont echter een veel groter oplossend vermogen en kleinere variantie dan het klassieke periodogram, bij lage frequenties.

Wil men uit een gegeven tijdreeks met lengte T , deze lage frequenties ($< 4/T$) goed detecteren, dan moet men een andere methode, zoals de exacte spectraalrepresentatie gebruiken (Zagalsky, 1967).

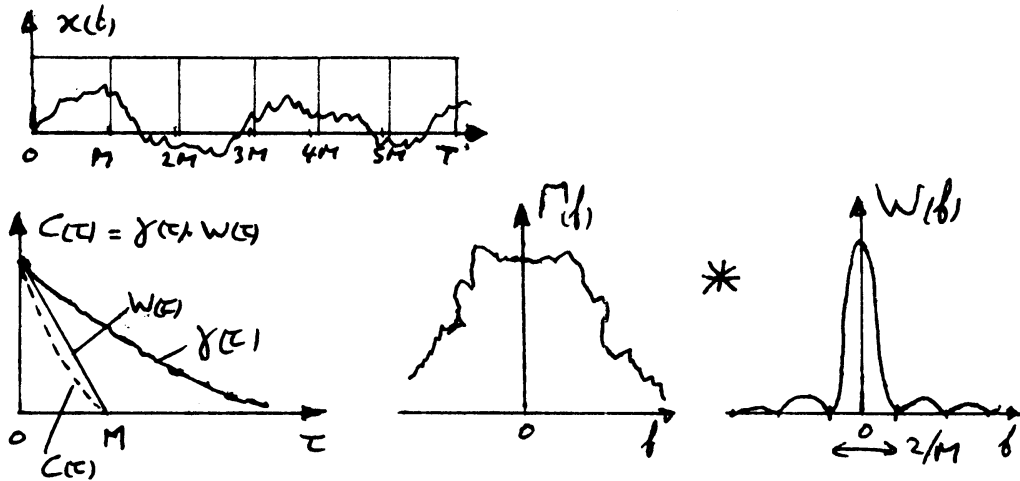
Hier wordt een oneindige Fouriertransformatie toegepast over een oneindige a.c.v.f. $\gamma(\tau)$ en er wordt geen spiegeling aangetroffen, die de bovengenoemde effecten veroorzaakt.

6. Vensters en filters.

Het effect van een tijdvenster in het spectrale domein wordt wel een filter genoemd. Vensters in het tijdsdomein of corresponderende filters in het spectrale domein worden gebruikt om de ruwe periodogramschatter $C(f)$ glad te maken. De filterwerking is een middeling in het spectrale domein. Een venster op de a.c.v.f. $\gamma(\tau)$ is bepaalde middeling van delen van de tijdreeks, b.v.: Bartlett toonde aan, dat de middeling van de spectra verkregen uit het in gelijke stukken knippen van de tijdreeks, gelijk was aan het toepassen van een driehoekig venster op de a.c.v.f. $\gamma(\tau)$ van lengte M gelijk de lengte van elk stuk.

Dit is in het frequentiedomein gelijk aan het toepassen van een $\frac{\sin M\omega}{\frac{1}{2}\omega}$ convolutie en in het spectrale domein het toepassen van het "smoothing" filter $(\frac{\sin \frac{1}{2}M\omega}{\frac{1}{2}\omega})^2 \frac{1}{M}$.

Door het in stukken knippen wordt de variatie gereduceerd met factor T/M .



Figuur 12. Het ondelen van $x(t)$ in k stukken, lengte M levert een $\Gamma(f)$ gezien door een filter $W(f)$ dat k x zoveel "afvlakt" als het filter $W(f)$ behorende bij de reeks $x(t)$ lengte T .
De variatie van $C(f) = \Gamma(f) * W(f)$ wordt een factor k kleiner, het oplossend vermogen echter ook, bovendien wordt de fout (bias) groter.

In formule wordt algemeen $\Gamma(f)$ door het spectrale venster $W(g)$ bekeken; gezien wordt de gewogen $\Gamma(f)$ nl. $\bar{C}(f)$,

$$\bar{C}(f) = \int_{-\infty}^{+\infty} W(g) \cdot \Gamma(f - g) dg,$$

waarbij $W(g)$ de Fouriergetransformeerde is van venster $w(\tau)$,

$$W(f) = \int_{-\infty}^{+\infty} w(\tau) e^{-j2\pi f \tau} d\tau.$$

$\bar{C}(f)$ is de Fouriertransform van $\bar{C}(\tau) = w(\tau) \cdot \gamma(\tau)$.

$$\bar{C}(f) = \int_{-\infty}^{+\infty} w(\tau) \cdot \gamma(\tau), e^{-j2\pi f \tau} d\tau.$$

Hieruit volgt, dat als $W(g)$ zich als δ functie gedraagt, of als $\Gamma(f)$ relatief vlak is over de bandbreedte B van $W(g)$, dat dan $\bar{C}(f)$ gelijk is aan $\Gamma(f)$, een monster van het spectrum van $X(t)$ over T . (zie voor de bandbreedte blz. 20).

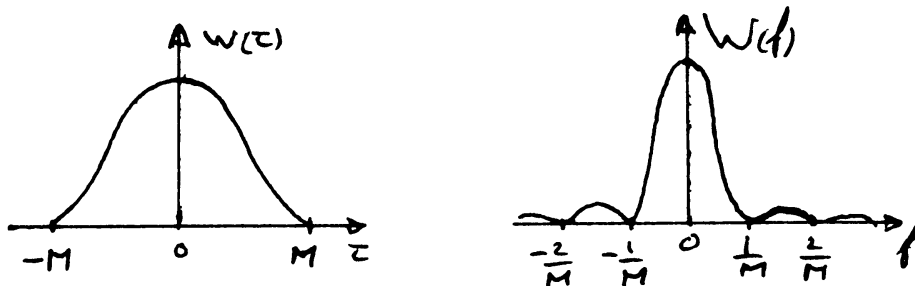
$$\bar{C}(f) = \int_{-\infty}^{+\infty} \delta(g) \Gamma(f-g) dg = \Gamma(f).$$

Om de filters te normaliseren wordt verondersteld, ook als $W(f)$ zich niet als δ functie gedraagt, dat

$$\int_{-\infty}^{+\infty} W(g) dg = 1.$$

De eigenschappen worden bekeken, waaraan het venster $w(\tau)$ en filter $W(f)$ moeten voldoen.

$w(\tau)$	F_0 →	$W(f)$
1) $w(0) = 1$		$\int_{-\infty}^{+\infty} W(f) df = w(0) \equiv 1$
2) $w(\tau) = w(-\tau)$		$W(f) = W(-f)$
3) $w(\tau) = 0, \tau \geq M < T$		$W(f)$ heeft een breedte van de hoofdlob $O\left(\frac{1}{M}\right)$



Figuur 13. Het convolutievenster $w(\tau)$ en spectraalvenster $W(f)$.

De meest gebruikte vensters zijn

het Tukeyvenster: $w_T(\tau) = \frac{1}{2}(1 + \cos \frac{\pi \tau}{M})$ voor $|\tau| \leq M$
 $= 0$ voor $|\tau| > M$

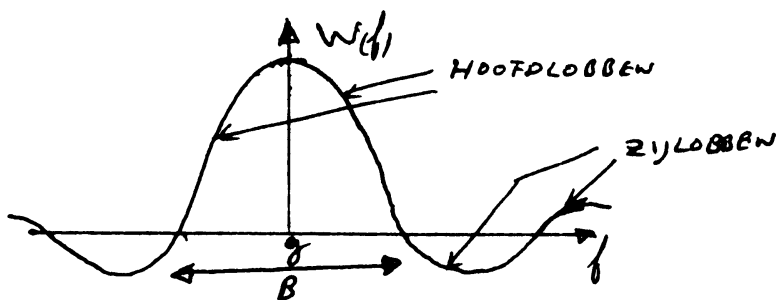
en het Parzenvenster: $w_p(\tau) = 1 - 6(\frac{\tau}{M})^2 + 6(\frac{|\tau|}{M})^3$ voor $|\tau| \leq M/2$
 $= 2(1 - (\frac{|\tau|}{M})^3)$ voor $\frac{M}{2} < |\tau| \leq M$
 $= 0$ voor $|\tau| > M$

$$W_T = M \left(\frac{\sin^2 \pi f M}{2 \pi f M} \right) \left(\frac{1}{1 - (2fM)^2} \right)$$

$$-\infty < f < +\infty$$

$$W_p = \frac{3}{4} M \left(\frac{\sin \pi f M / 2}{\pi f M / 2} \right)^4$$

Het voordeel van het Tukey of Parzenfilter is de aanwezigheid van kleine zijlobben en de daaruit voortvloeiende kleine lek (bijdrage van andere frequenties dan binnen de bandbreedte).



Figuur 14. De zijlobben van een frequentievenster $W(f)$ geven foutieve bijdragen van spectrale energie buiten de bandbreedte aan frequentie g .

Het doel van een venster is om de variantie van het samplespectrum te verkleinen met factor $\sim \frac{T}{M}$. Echter ontstaat bij niet gladde spectra een fout t.o.v. het spectrum I . Deze fout wordt de bias genoemd ($b(f)$).

$$b(f) = E(\bar{C}(f)) - \Gamma(f),$$

$$\text{met } E(\bar{C}(f)) = \int_{-\infty}^{+\infty} W(g) \cdot E(c(f-g)) dg \approx \int_{-\infty}^{+\infty} W(g) \Gamma(f-g) dg.$$

Ideaal is een $w(f)$, die zich als δ functie gedraagt.

$$\text{dan } E(\bar{C}(f)) = \int_{-\infty}^{+\infty} \delta(g) \Gamma(f-g) dg = \Gamma(f).$$

Dus uit grote M is een zuiver (bias = 0) spectrum te bepalen voor de reeks $x(t)$, $t = (0 \dots T)$. Om de variantie klein te houden moet T een groot (> 10) aantal malen M bedragen.

De fout, die gemaakt wordt bij een Tukey of Parzenfilter, kan worden afgeschat uit de tweede afgeleide van $\Gamma(f)$, die we echter niet kennen, b.v. (Jenkins en Watts, 1968, blz. 247).

$$b_T(f) = \frac{0.036}{M^2} \Gamma^{(2)}(f) + O\left(\frac{1}{M^4}\right) \quad \text{Tukey (15.2)}$$

$$b_P(f) = \frac{0.152}{M^2} \Gamma^{(2)}(f) + O\left(\frac{1}{M^3}\right) \quad \text{Parzen (15.3)}$$

Het spectrum Γ wordt dus onderschat bij een piek (\rightarrow lagere piek) en overschat bij een dal.

Uit (15.3) of (15.4) kan uit twee stabiele (zie verder) schatters C_1 en C_2 met verschillende M_1 en M_2 een schatter voor $M(f)$ worden gevonden met kleinere bias.

$$\Gamma(f) \approx C_1(f) + (C_1(f) - C_2(f)) / \left(\left(\frac{M_1}{M_2} \right)^2 - 1 \right) \text{ als } M_1 > M_2.$$

De variantie van een gefilterd spectrum $C(f)$:

$$\text{Var}(C(f)) = \Gamma^2(f) \cdot \frac{I}{T} \text{ met } I = \int_{-\infty}^{+\infty} W^2(g) dg = \int_{-\infty}^{+\infty} W^2(\tau) d\tau \quad (15.4)$$

De bandbreedte van een filter B bepaalt het oplossend vermogen van het samplespectrum C(f).

$$B = \frac{1}{I} = O\left(\frac{1}{M}\right) \quad (15.5)$$

Voor het Tukeyvenster geldt $\frac{I}{T} = 0,75$; $B = \frac{1,33}{M}$.

Voor het Parzenvenster geldt $\frac{I}{T} = 0,54$; $B = \frac{1,86}{M}$.

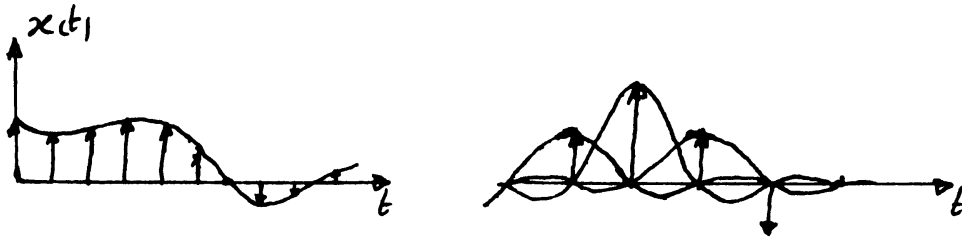
Deze twee venster zijn volkomen vergelijkbaar, zij het dat om dezelfde resultaten te verkrijgen voor Parzen een iets grotere M moet worden gekozen, zodat de bandbreedte en variantieverhouding gelijk worden. Algemeen geldt voor deze vensters

$$\text{Variantieverhouding} \times \text{bandbreedte} = \text{konstant} = \frac{1}{T}.$$

7. Aliasing.

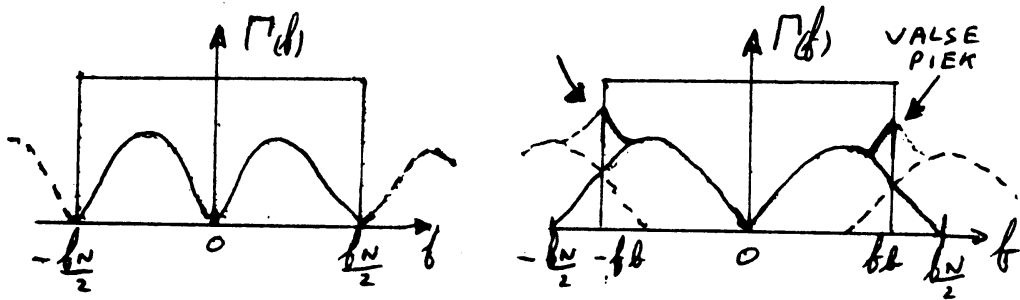
Bij het bemonsteren van de variabelen X(t) moet gezorgd worden dat aan het Nyquistkriterium wordt voldaan.

Dit luidt: De bemonsteringsfrequentie f_b moet minstens twee maal zo groot zijn als de hoogste frequentie, die in het signaal x(t) voorkomt. Ook is dan uit de bemonsterde reeks weer de continue reeks x(t) exact te konstrueren.



Figuur 15. De bemonsterde reeks en rekonstructie m.b.v. $\frac{\sin f_b t}{t}$ als $f_b > 2 f_{\max} = f_N$ (De Nyquistfrequentie).

Als hier echter niet aan voldaan wordt, treedt aliasing of optelling van een deel van de spiegeling t.o.v. de bemonsteringsfrequentie op in het spectrale domein.



Figuur 16. Bij bemonsteren wordt een rechthoekig bemonsteringsvenster met het spectrum vermenigvuldigd van breedte $2f_b$. Het spectrum ziet zichzelf altijd gespiegeld t.o.v. de bemonsteringsfrequentie. Bij foute keus $f_b < f_N$ wordt het gespiegelde deel bij het oorspronkelijke spectrum ongeteld.

8. Referenties bij Inleiding en deel I.

1. Schuster, 1898
2. Blackman en Tukey, 1958 - The Measurements of Power Spectra,
New York, Dover, 1958.
3. Radar en Gold, 1965 - Digital Filter Design Techniques in the
frequency domain, Proc. IEEE, V55, 1965.
4. Cooley, Lewis, Welch, 1967 - Historical Notes on FFT,
IEEE, Vsl AU-15, No. 2, P76, P79, 1967.
5. A. v.d. Bos, 1970 - An alternative interpretation of MESA
IEEE, trans. inf. th., juli 1971.
6. Burg, J.P., 1967 - Maximum Entropy Spectral Analysis,
Meeting Soc. Expl. Geoph. science,
Oklahoma, 1967.
7. Ulrich, 1972 - Maximum Entropy Power Spectrum of
Truncated sines, J. of Geoph. Res.,
Vol. 77, 1972.
8. Zagalsky, 1967 - Exact Spectral Representation of Truncated
data, Proc. letters of IEEE, jan. 1967.
9. Schikedanz en Bowen, 1975 - Computation of climatological power
spectra using VRL, Fourth Conf. on Prob.
St. in Atm. Sc., Nov. 1975.
10. Jenkinson - MET 013, Branch MEMO No. 57, 1977.
11. Jenkins en Watts, 1968 - Spectral Analysis and its applications,
Holden Day, San Francisco, 1968.
12. Toman, K., 1965 - The Spectral shifts of truncated sines,
Journal Geo. Res., Vol. 70, 1965.
13. Chen en Stegen, 1974 - Experiments with Maximum Entropy Power
Spectra of sines, J. Geo. Res., Vol. 79,
1974.

Deel II: Bespreking en vergelijking van verschillende methoden voor het berekenen van spectra.

Achtereenvolgend worden nu de methoden behandeld; waar ze op gebaseerd zijn; wat hun toepassingsgebied is en hoe de statistische eigenschappen afgeschat kunnen worden. Ook de benodigde rekentijd wordt bekeken.

Na het behandelen van de methoden volgt nog een samenvatting, wat betreft toepassingsmogelijkheden.

De methoden worden vergeleken. Conclusies worden getrokken en aanbevelingen gedaan.

9. De periodogram-schatter.

De ruwe periodogram-schatter kan zowel direct als indirect uit de eindige tijdreeks worden bepaald.

De directe methode: $C(f)$ wordt bepaald uit directe Fourier-transformatie van de eindige tijdreeks $x(t)$, $-\frac{T}{2} < t < \frac{T}{2}$:

$$C(f) = \frac{1}{T} \left[\int_{-\frac{T}{2}}^{+\frac{T}{2}} x(t) e^{-2\pi jft} dt \right]^2 .$$

Dit gebeurt voor grotere discrete tijdreeks $x(n)$, $n = 0 \dots N$, snel door het door Cooley en Tukey (1965) ontwikkelde Fast Fourier algoritme (F.F.T.). Als één bewerking wordt gedefinieerd als één optelling en één vermenigvuldiging, kost de F.F.T. methode tijd voor $2N \log N$ bewerkingen, tegenover $2N^2$ voor ontbinding van $x(t)$ in sinussen en cosinussen, zonder speciaal algoritme.

Bij de indirecte methode wordt $C(f)$ verkregen door Fourier-transformatie van de a.c.v.f. $c(\tau)$, verkregen uit de eindige tijdreeks $x(t)$ met $0 < t < T$

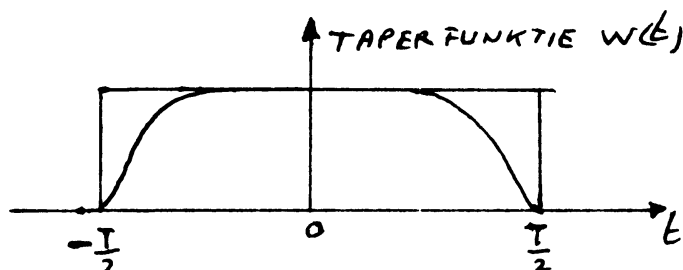
$$C(f) = 2 \int_0^T c(\tau) e^{-2\pi jf\tau} d\tau .$$

De $c(\tau)$ is symmetrisch om nul en kan ook met een F.F.T. algoritme snel uit $x(n)$, met $n = 0 \dots N$ worden bepaald.

Beide methoden zijn gelijkwaardig (zie vgl. 1.1).

Het eenvoudigste is om directe Fourier-transformatie toe te passen. Dit kost minder computertijd, dan de indirecte methode. De autocovariantiefunctie $C(\tau)$ wordt echter niet verkregen. Deze kan belangrijke informatie verschaffen over de lengte van het eventueel toe te passen venster $W(\tau)$. Door de eindige Fourier-transformatie treedt een periodiciteit op gelijk aan de lengte van de transformatie (T) en bovendien worden hogere harmonischen in $C(f)$ geïntroduceerd door de discontinuïteit in de oneindige, steeds gespiegelde te transformeren functie (§ 5). De hogere harmonischen zijn de zijlobben van het spectraalvenster $W(f)$.

Golflengten van orde T of groter geven als gelijkspanningsdeel bijdragen tot fouten in het periodogram. Een verbetering kan worden verkregen door een taper toe te passen op de tijdreeks. De discontinuïteit aan het begin en einde van de reeks, die de hogere harmonischen veroorzaakt wordt zo gedeeltelijk te niet gedaan. Een bias is hiervan echter het gevolg, daar in feite een venster wordt toegepast op de a.c.v.f.



Figuur 17. Het toepassen van een anders dan rechthoekig venster over de tijdreeks $x(t)$ in het tijdsdomein (cosinustaper) om de discontinuïteit van de data te verminderen.

Tapering is van belang, bij toepassen van het periodogram op niet te lange tijdreeksen (Koopmans, 1974). Het taperen komt neer op het toepassen van een venster $W(\tau)$ op $\gamma(\tau)$, gelijk aan de autocovariantiefunctie van de taperfunctie.

De ruwe periodogram-schatter is zuiver (unbiased) als de discontinuïteit t.a.v. de eindige Fourier-transformatie buiten beschouwing wordt gelaten. Dit is juist bij voldoende grote tijdreekslengte, t.o.v. de laagste frequenties, waarbij de bijdragen van de hogere harmonischen te verwaarlozen zijn t.o.v. het zuivere spectrum $\Gamma(f)$.

De dan zuiver veronderstelde schatter heeft echter een variantie, die aanzienlijk is en onafhankelijk van de meetlengte T.

De variantie van C(f) is gelijk aan $\Gamma^2(f)$ of

$$\frac{\text{Var. (C(f))}}{\Gamma^2(f)} = 1 .$$

Dit geldt bij benadering voor niet te kleine T (Jenkins en Watts, blz. 250). De invloed van de zijlobben van het spectrale venster is dan niet te groot en de frequentie bijdragen op afstand $\frac{2}{T}$ zijn dan onafhankelijk te stellen. Het spectrale venster is een Bartlett venster.

Gemeten wordt in feite:

$$\bar{C}(f) = \int_{-\infty}^{+\infty} W(g - f) \Gamma(f) dg \text{ met } ,$$

$$W(f) = W_B(f) = T \left(\frac{\sin \pi f T}{\pi f T} \right)^2 .$$

Dus een onzuiver spectrum met bandbreedte $\frac{2}{T}$.

Als $T \rightarrow \infty$ gedraagt W(f) zich als δ functie en is het ruwe periodogram zuiver (unbiased).

Een speciaal periodogram is het Jenkinson periodogram (Jenkinson, 1977).

Zowel de frequenties als de golflengte zijn logaritmisch verdeeld.

Bij Jenkinson komen de golflengten L_m en frequenties F_m op logaritmisch gelijke afstanden te liggen:

$$L_m = 2 \cdot e^{\lambda m} \quad \text{en} \quad F_m = \frac{1}{L_m} = 0,5 \cdot e^{-\lambda m} .$$

λ heet de "tuning", die het spectrum over de frequentieband kan verschuiven van $f = 0,5$ tot $f = 0$ voor $\lambda = 0$ tot ∞ .

Het voordeel van deze logaritmische verdeling is het hogere oplossend vermogen bij de lage frequenties. Relatief worden daar meer frequenties meegenomen.

10. De venstermethoden.

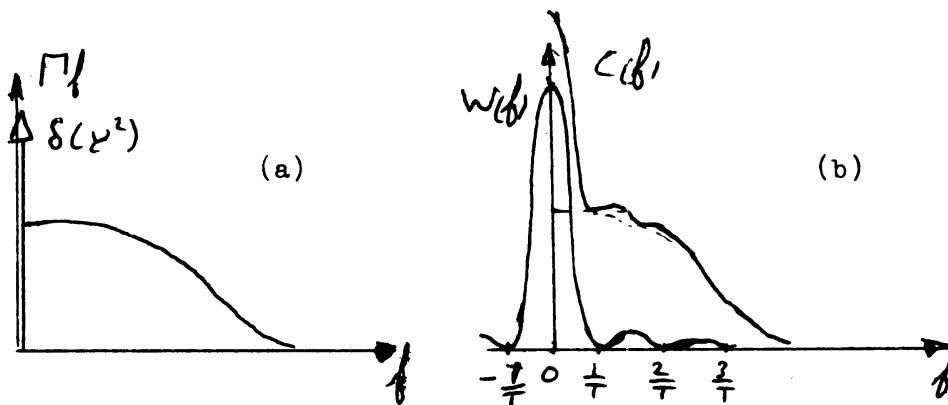
De ruwe periodogram-schatter dient als basis voor de venstermethoden. Met behulp van een impliciet of expliciet toegepast venster, wordt de variantie van deze schatter gereduceerd. Dit kan op de volgende manieren geschieden:

1. Het opdelen van de tijdreeks in k gelijke delen en middeling van de hieruit verkregen k ruwe periodogram-schatters levert een reductie van de variantie van het spectrum met factor k .
Dit is het impliciet toepassen van een Bartlett venster (lengte $\frac{T}{k}$).
2. Berekening van de a.c.v.f. $C_a(\tau)$ uit $x(t)$ met $t = 0 \dots T$ en het toepassen van een goed venster in het τ -domein met lengte M , zoals het Tukey of Parzenvenster. Dit levert een bandbreedte $B = O(\frac{1}{M})$ en een variantiereductie $O(\frac{T}{M})$.
Zie voor precieze waarden, Jenkins en Watts, 1968, blz. 252 - 244.
Een goed filter bezit kleine zijlobben.
3. Het glad maken van de over de hele tijdreeks berekende ruwe periodogram-schatter, met behulp van een spectraal venster. Klassiek is dit venster (deze "smoothing") onafhankelijk van de frequentie. Daar echter fouten ontstaan (bias), bij pieken door dit glad maken, lijkt het een goede methode om dit filter af te laten hangen van de vorm van het ruwe periodogram.
Bij glad lijkende delen kan een breed spectraal-venster worden toegepast, om een kleine variantie te krijgen met kleine bias. Er wordt dan met behulp van veel naburige discrete frequenties gemiddeld.
Bij pieken wordt de bandbreedte van het spectrum in overeenstemming gebracht met de breedte van de piek. Er wordt over weinig punten gemiddeld in het spectrum.
De pieken worden met grotere variantie, maar met kleinere bandbreedte bepaald, dan de delen met kleine tweede afgeleide (de gladde delen).
De variantieverhouding \times bandbreedte is konstant ($1/T$). Deze methode kan misschien geautomatiseerd worden en een veel beter spectrum kan worden verkregen, uit het beschikbare ruwe periodogram, dan met de klassieke methoden met een star frequentie onafhankelijk spectraal venster. Op het oog smoothen is vergelijkbaar met deze methode en vermoedelijk beter dan enige vaste venster methode.

Jenkins en Watts (1968) bevelen een techniek aan, waarbij eerste met een spectraal-venster met grote bandbreedte het spectrum bekeken wordt. Het spectrum is dan glad en stabiel, maar vertoont echter bias, die zeer groot kan zijn bij aanwezigheid van pieken. Vervolgens wordt het spectraal-venster smaller gemaakt. Er komen meer details naar voren, echter ook instabiliteiten of valse pieken. De techniek heet "window closing" en geeft inzicht in de vorm van het spectrum.

We geven nu nog enkele opmerkingen over de fouten, die gemaakt kunnen worden bij de aanwezigheid in het periodogram van een piek op $f = 0$ ten gevolge van het gemiddelde van $x(t)$, en andere pieken.

Als de verwachtingswaarde van $X(t)$, nl. $E(X(t))$ ongelijk nul is, leidt dit in het echte spectrum $\Gamma(f)$ van $X(t)$ tot een discontinuïteit bij frequentie nul. $E(X(t)) = \mu$ wordt weergegeven als δ functie (figuur 18a).



Figuur 18. Het ware spectrum Γ ~~en~~ het ruwe periodogram C , met spectraal-venster W .

Bij het ruwe periodogram $C(f)$, verkregen uit eindige meettijd T van $x(t)$, wordt dit gemiddelde spectraal getransformeerd tot $W(f)$ en dus gelijk aan het spectrale venster van het ruwe periodogram (figuur 18b). Dit heeft eindige hoogte en zijlobben.

Het gemiddelde \bar{x} van $x(t)$ over de meettijd T vormt een blokfunctie in de tijd. De energie-inhoud is gelijk aan

$$\int_0^T \bar{x}^2 dt = T\bar{x}^2 .$$

Deze energie wordt verdeeld in het spectrale domein door de functie $W(f)$.

Het Parsevaltheorema zegt dat

$$\int_0^T \bar{x}^2 dt = \int_{-\infty}^{+\infty} F_0^2(\bar{x}) df = \int_{-\infty}^{+\infty} W(f) df ,$$

zodat

$$\int_{-\infty}^{+\infty} W(f) df = T \bar{x}^2 .$$

Eenvoudig is nu te zien dat

$$W(f) = \left(\frac{\sin \pi T \cdot f}{\pi T f} \right)^2 \cdot T^2 \cdot \bar{x}^2 ,$$

omdat

$$\int_{-\infty}^{+\infty} T \cdot \left(\frac{\sin \pi T f}{\pi T f} \right)^2 df \equiv 1 \quad (\text{zie blz. } 13).$$

Voor de nulfrequentie volgt

$$W(0) = T^2 \bar{x}^2 .$$

De 1e harmonische heeft een relatieve top op $f = \frac{3}{2T}$ van 4,4%

De 2e harmonische heeft een relatieve top op $f = \frac{5}{2T}$ van 1,6%

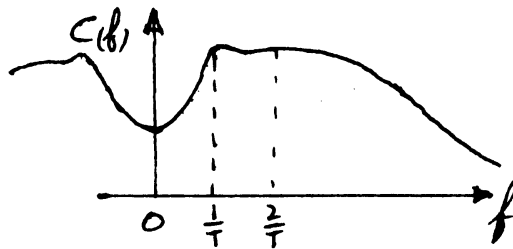
De 3e harmonische heeft een relatieve top op $f = \frac{7}{2T}$ van 0,8%

Dit ten opzichte van de top van $W(0)$.

Dit lijken kleine bedragen maar als bijvoorbeeld de piek 20 x boven het spectrum uitsteekt, wat een redelijke waarde is, dan zijn de relatieve fouten t.o.v. het spectrum resp. 88%, 32% en 16%.

Dit verschijnsel kan worden voorkomen als de verwachtingswaarde $E(x(t))$ of \bar{x} voor $T^* \gg T$ van de reeks wordt afgetrokken, waarna van $x(t) - E(x(t))$ het periodogram wordt bepaald, wat geen bias meer vertoont bij de nul-frequentie.

Vaak is slechts \bar{x} over T bekend. Als dit van de reeks $x(t)$ wordt afgetrokken, wordt uit het spectrum C , $W(0)$ weggehaald. Dit kan een dip in het spectrum van breedte ongeveer $\frac{1}{T}$ veroorzaken. Deze is echter veel minder diep, dan de $W(0)$ functie hoog is. Als na het verwijderen van \bar{x} over T gefilterd wordt met behulp van een b.v. 10 x zo breed venster als $W(f)$ zal de dip praktisch verdwijnen, echter met een beetje bias rond nul.



Figuur 19. Bij verwijderen van \bar{x} ontstaat een dip in $C(f)$, die niet altijd naar nul hoeft te gaan, van breedte $\frac{1}{T}$.

Een nog ernstiger fout wordt gemaakt als het gemiddelde over T niet wordt verwijderd en men een spectraal-venster toepast om te smoothen met grotere bandbreedte dan $\frac{2}{T}$. De energie in het gemiddelde zal dan uitgesmeerd worden over een evenredig breder spectraal-gebied.

De piek bij $f = 0$ wordt minder hoog. In het spectrum kunnen we de zijlobben nu terug vinden tot over grote f , evenals bij figuur 13 met $M = \bar{T}$.

Aanbeveling:

1. Bereken indien mogelijk $E(x(t)) = \frac{1}{T^*} \int_0^{T^*} x(t) dt$ met $T^* \gg T$.
Pas nu de F.F.T. routine toe, met of zonder venster op de reeks $x(t) - E(x(t))$.
2. Of, als $E(x(t))$ niet bekend is, neem dan $\bar{x} = \frac{1}{T} \int_0^T x(t) dt$.
Het ruwe periodogram vertoont een dip bij $f = 0$. Verwijder nu deze foutieve dip in $C(f)$ door extrapolatie van $C(f)$ naar $f = 0$. Hierna smoothen op het oog. Als men een spectraal-venster gebruikt zal de dip verminderen.

Hetzelfde kan gebeuren met een "deterministische" scherpe piek in het spectrum. Het beste kan weer deze sinus vooraf verwijderd worden. Als dit niet kan, kunnen de zijlobben met het oog worden opgespoord en verwijderd worden.

11. De variabele recordlength-methode.

Deze methode is een voorstel van Schickedanz en Bowen (1975) en wil een hoger oplossend vermogen verkrijgen bij de lage frequenties, door meer lagere frequenties mee te nemen.

Dit gebeurt door het herhaald toepassen van een F.F.T. routine op een verkorte en verlengde tijdreeks.

Bij verkorting worden wederzijds één of meer samples weggelaten.

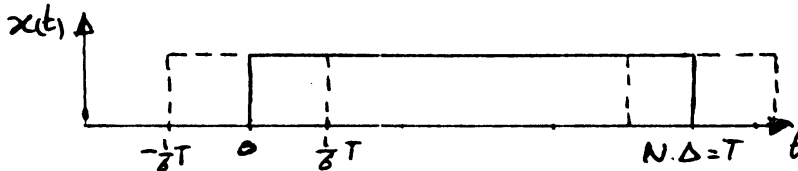
Bij verlenging wordt wederzijds een aantal malen de gemiddelde waarde van de tijdreeks $E(x(t))$ toegevoegd.

De spectra, die nu uit deze reeksen worden verkregen, worden genormeerd met de energie, die aanwezig is in de respectievelijke spectra. Vervolgens wordt het VRL spectrum verkregen door middeling van de genormeerde spectra.

Dit VRL-spectrum is veel gedetailleerder dan het periodogram, vooral bij lage frequenties. Eventueel kan nog eenlichte smoothing worden toegepast om de variantie te verkleinen.

Het criterium, dat aangeeft hoeveel keer de F.F.T. routine moet worden toegepast op de verkorte of verlengde reeks is afhankelijk van hoe fijn de gewenste verdeling is.

Schickedanz en Bowen benutten elke mogelijk te berekenen frequentie. Het lege gebied tussen het 1e en 2e golfgetal bij het oorspronkelijke F.F.T. spectrum is dan opgevuld als de reeks wederzijds met $1/6$ zowel vergroot als verkleind wordt. Er moeten dan dus $(\frac{1}{3}N)$ F.F.T. routines worden uitgevoerd, zie figuur 20.



Figuur 20. Het berekenen van het ruwe spectrum van $x(t)$ geschiedt bij de VRL-techniek volgens Schickedanz en Bowen $\frac{1}{3}N$ maal over een verkorte en verlengde tijdreeks.

Opmerkingen over VRL.

In feite wordt niets anders gedaan dan inter- en extrapolatie tussen de golfgetallen. Het spectrum wordt gedetailleerder. Toch wordt meer informatie uit het spectrum gehaald bij de VRL-techniek dan conventioneel, daar de stelling, dat spectrale-schatters op frequentie-afstand van de filterbandbreedte $O(\frac{1}{T})$ ongeveer onafhankelijk zijn, niet inhoudt dat tussenliggende schatters uit deze conventionele kunnen worden opgebouwd. Dit komt door de vorm van het spectrale filter.

Hier is dat een $(\frac{\sin T \omega}{\omega})^2$ functie. Deze is ongelijk aan de vorm van het ideale bandpassfilter, waarbij inderdaad geen nieuwe informatie zou worden verkregen bij de VRL-techniek. De niet ideale vorm van het spectrale filter maakt de VRL-techniek mogelijk, tevens ook geeft deze zin aan de Jenkinson-methode (Jenkinson, 1977).

Bedacht moet worden, dat bij vergroting van de reekslengte de bandbreedte van het filter gelijk blijft aan de oorspronkelijke, ook de invloed van de discontinuïteit in $x(t)$ aan de uiteinden blijft dezelfde. Bij verkleining wordt echter de bandbreedte omgekeerd evenredig groter met de reekslengte, bovendien wordt de invloed van de discontinuïteit relatief groter, wat de hogere harmonischen en frequentieverschuivingen bij lage frequenties versterkt. Dit is mijn voornaamste kritiek op de VRL.

Beter kan direct een andere frequentie-verdeling zoals, die van Jenkinson worden toegepast.

De VRL kost bovendien erg veel rekentijd n.l. $\frac{2}{3}N^2 \ln N$.

Het voordeel is, dat VRL eenvoudig kan worden toegepast met alleen een F.F.T. routine aanwezig.

12. De maximale entropie methode (MEM).

Het MEM spectrum van een stationair random, uniform bemonsterd, proces is het spectrum dat resulteert uit het maximaliseren van de entropie van het proces. Dit blijkt uit de hierna volgende redenering.

Informatie en entropie. We beginnen met een statistische definitie van informatie.

Beschouw een situatie waarin M verschillende gebeurtenissen m_i plaats kunnen vinden, met waarschijnlijkheden van optreden P_i . Als de voorwaarde bestaat dat alle P_i gelijk zijn, dan kan er geen speciale informatie van het systeem worden verkregen.

Als we echter één speciale P_i kunnen specificeren krijgen we een zekere informatie van het systeem. Het is duidelijk dat de waarschijnlijkheid van optreden van een gebeurtenis verbonden is met informatie. Als de gebeurtenis waarvoor we een waarschijnlijkheid hebben gespecificeerd, onverwacht is, dan geeft dit veel informatie over het systeem. We schrijven daarom het verband tussen de informatie en waarschijnlijkheid als

$$I = - \sum \ln \frac{1}{P_i} .$$

Het nut van de logaritme is gerechtvaardigd, als we de informatie als een additieve grootte zien.

Laten we nu veronderstellen, dat we het systeem voor een lange tijd T bekijken. Als T erg groot is kunnen we $P_i \cdot T$ gebeurtenissen m_i verwachten in dit tijdsinterval.

De totale informatie van het systeem wordt dan

$$I \text{ totaal} = \sum_{i=1}^M P_i T \ln \frac{1}{P_i} = T \sum_{i=1}^M P_i \ln \frac{1}{P_i} .$$

De gemiddelde informatie per tijdsinterval wordt voorgesteld door H en wordt de entropie genoemd. Dus (Shannon, 1948),

$$H = \frac{I \text{ totaal}}{T} = - \sum_{i=1}^M P_i \ln P_i .$$

Het is duidelijk uit de bovenstaande discussie over de relatie tussen informatie en waarschijnlijkheid, dat entropie een maat is voor onzekerheid beschreven door een set van waarschijnlijkheden. De entropie is nul voor een systeem waar alle P_i nul zijn behalve één, welke één is. In dat geval is het systeem volledig bepaald en bestaat er geen onzekerheid (deterministisch signaal). In alle andere gevallen is de entropie positief ($P_i \leq 1$).

Hoe kan men het maximaliseren van de entropie een spectrale dichtheidsfunctie worden verkregen. De problemen die optreden met de conventionele methoden voor het bepalen van een spectrum zijn te wijten aan veronderstellingen die gemaakt worden over de gegevens buiten de parameterruimten.

Dit periodiek voortzetten of nul stellen van deze gegevens buiten het bekende interval resulteert vaak in onzuivere schatters. Wat nu vereist is, is een benadering die allereerst consistent is met de aanwezige gegevens en ten tweede de waarschijnlijkheidsverdeling hiervan schat, zonder gebruik te maken van gegevens buiten het bekende interval.

Jaynes (1963, 1968) formuleerde een buitengewoon succesvol principe over het verband tussen entropie en onzekerheid.

Dit luidt: de waarschijnlijkheidsverdeling, die de aanwezige informatie beschrijft maar maximaal niet verbonden is met de niet aanwezige informatie, is die met een maximum aan entropie.

We zien direct dat het gebruik van dit principe al de foutieve veronderstellingen over de niet aanwezige gegevens elimineert, die gebruikt worden in conventionele spectraal-analyse.

De methode om de maximale entropie waarschijnlijkheidsverdeling te bepalen is nu als volgt (Jaynes, 1968):

Het proces x_t dat wordt beschouwd kan waarden x_1, x_2, \dots, x_n aannemen. We veronderstellen dat de informatie over x_t beschikbaar is in de vorm van gemiddelde waarden $E(f_1(x_t)), \dots, E(f_m(x_t))$ van functies $f_1(x_t), \dots, f_m(x_t)$, $m < n$. De waarschijnlijkheidsverdeling $P_i = p(x_t = x_i)$ die consistent is met de aanwezige informatie, maar maximaal vrij van veronderstellingen is die, welke de entropie H maximaliseert.

$$H = -\sum_i P_i \ln P_i .$$

Dit onder de nevenvoorwaarden:

$$\sum_i P_i = 1,$$

$$\text{en } \sum_i P_i f_k(x_i) = E(f_k(x_t)), k = 1, 2, \dots, m.$$

De bekende oplossing voor dit variatie-probleem is:

$$P_i = \frac{1}{Z(\lambda_1 \dots \lambda_m)} \exp (\lambda_1 f_1(x_i) + \dots + \lambda_m f_m(x_i)) .$$

met $Z(\lambda_1 \dots \lambda_m) = \sum_i \exp(\lambda_1 f_1(x_i) + \dots + \lambda_m f_m(x_i))$.

De Lagrange parameters λ_k worden bepaald uit de nevenvoorwaarden, wat leidt tot de relaties:

$$E(f_k(x_t)) = \frac{\partial}{\partial \lambda_k} \log Z(\lambda_1 \dots \lambda_m).$$

Jaynes (1968) liet bovendien zien, dat de waarschijnlijkheidsverdeling die de entropie maximaliseert, numeriek gelijk is aan de frequentie-verdeling die op het grootste aantal manieren kan worden gerealiseerd. Deze frequentie-verdeling is de verdeling van $u_i = m_i/M$, met m_i het aantal keren dat x_i plaatsvindt in M realisaties van het proces x_t . Om het concept de maximale entropie op de spectraal-analyse toe te passen beginnen we met de relatie tussen de entropie (beter de entropiegraad van een oneindig proces) en de spectrale dichtheid $S(f)$ van een stationair Gaussisch proces,

$$H = \frac{1}{4f_N} \int_{-f_N}^{+f_N} \log S(f) df, \quad (1)$$

met f_N de Nyquistfrequentie. De afleiding van (1) staat gedetailleerd in Smylie et al (1973). In termen van autokorrelatie $\phi(k)$ van het proces kan (1) geschreven worden als:

$$H = \frac{1}{4f_N} \int_{-f_N}^{+f_N} \log \left(\sum_{-\infty}^{+\infty} \phi(k) \exp(-i2\pi f k \Delta t) \right) df. \quad (2)$$

met Δt het uniforme bemonsteringsinterval.

Deze vergelijking (2), gemaximaliseerd naar de onbekende $\phi(k)$ met nevenvoorwaarde dat $S(f)$ ook consistent moet zijn met de bekende autokorrelaties $\phi(0), \dots, \phi(M-1)$ levert de MEM spectraal-schatter.

Deze schatter is consistent met de aanwezige informatie maar bezit maximale onzekerheid t.o.v. de niet aanwezige informatie.

De variatie rekening leidt tot de expressie voor de MEM spectrale dichtheid voor een reëel lineair proces x_t .

$$P_E(f) = \frac{P_M}{f_N \left[1 + \sum_{j=1}^{M-1} \gamma_j \exp(-2 \pi F_j \Delta t) \right]^2} \quad (\text{Smylie et al, 1973}) \quad (3)$$

In (3) is P_M een konstante en γ_j zijn de voorspellingsfouten, die beide uit de gegevens bepaald worden. Ongetwijfeld is het belangrijkste tekortkomen van de MEM schatter geweest, het gebrek aan een kwantitatieve methode om de lengte van het voorspellingsfoutenfilter γ_j in (3) te bepalen. Werk van Akaike (1970) over de bepaling van de orde van autoregressieve processen heeft dit probleem uit de wereld geholpen. Om dit te laten zien moeten we eerst het verband laten zien tussen MEM en de autoregressieve representatie van een stochastisch proces.

We zijn geïnteresseerd in de klasse van discrete normaal verdeelde random processen, die bovendien lineair zijn.

Het proces x_t dat hiertoe behoort kunnen we benaderen door een eindig aantal parameters x_i en een ongecorreleerde ruïsterm a_t ,

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \dots + \alpha_p x_{t-p} + a_t \quad (4)$$

Deze vergelijking (4) is een discrete vertegenwoordiging van een autoregressief (AR) proces van de orde p . Voor stationariteit moeten de wortels van het polynoom $1 - \alpha_1 z - \dots - \alpha_p z^p$ binnen de eenheidscirkel liggen.

Eindige orde AR processen zijn van groot belang in de voorstelling van fysisch belangrijke lineaire random processen. Bijvoorbeeld, een tweede orde continue AR proces kan worden beschreven als

$$k_2 \frac{d^2 x(t)}{dt^2} + k_1 \frac{dx(t)}{dt} + x(t) = a(t) \quad .$$

Deze vergelijking beschrijft het gedrag van een gedempt lineair tweede orde systeem, een elektrisch circuit bijvoorbeeld (oscillator), die aangestoten wordt door de random impulsen $a(t)$. Een AR model wordt ook wel een meer-polen-model (all pole) genoemd. De ongecorreleerde (witte) ruïsterm a_t heet wel de innovatie van het AR proces.

De overeenkomst tussen een AR proces en MEM.

We volgen nu de ontwikkeling van Van den Bos (1971), die het verband aangaf tussen de MEM spectraal analyse en de AR representatie van een random proces. Deze discussie leidt tot de methoden om de AR parameters van een gegeven proces te bepalen en dus hieruit het spectrum. Smylie et al (1973) liet een uitdrukking zien voor de entropie van een discreet stationair Gaussisch proces x_t , $t = 1, 2, \dots, M + 1$, welke equivalent is aan (1):

$$H = \frac{1}{2} \log (\det(C(M))) , \quad (5)$$

met $C(M)$ de semipositief definitieve Toeplitz autocovariantie matrix van het proces x_t (Gray, 1972).

$$C(M) = \begin{bmatrix} \rho(0) & \rho(1) & \dots & \rho(M) \\ \rho(1) & \rho(0) & \dots & \rho(M-1) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(M) & \rho(M-1) & \dots & \rho(0) \end{bmatrix} \quad (6)$$

met $\rho(N)$ de autocovariantie. Veronderstel dat de eerste $M + 1$ autocovarianties $\rho(0)$, $\rho(1)$, \dots , $\rho(M)$ exact bekend zijn.

Het idee achter de MEM is nu, de onbekende autocovariantie coëfficiënten $\rho(M + 1)$, $\rho(M + 2)$, etc., zo te kiezen, dat de entropie van het proces maximaal is volgens (5) bij iedere stap.

Uit (5) volgt dan, dat $\rho(M + 1)$ wordt bepaald door $\det (C(M + 1))$ te maximaliseren naar $\rho(M + 1)$. Het feit dat deze determinant een eenduidig maximum heeft als functie van $\rho(M + 1)$ volgt uit het semipositief definitief zijn van $C(M + 1)$ door de produkt regel van de differentiatie toe te passen. Als eenmaal $\rho(M + 1)$ is bepaald, kan $\rho(M + 2)$ bepaald worden door $\rho(M + 1)$ in $C(M + 2)$ te substituëren.

Deze procedure wordt herhaald voor de andere coëfficiënten.

De oplossing van de maximalisatie van $\det (C(M + 1))$ naar $\rho(M + 1)$ is

$$\det \begin{bmatrix} \rho(1) & \rho(0) & \dots & \rho(M-1) \\ \rho(2) & \rho(1) & \dots & \rho(M-2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(M+1) & \rho(M) & \dots & \rho(1) \end{bmatrix} = 0 \quad (7)$$

We gaan nu terug naar het AR proces (4) met orde M.

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \dots + \alpha_M x_{t-M} + a_t \quad (8)$$

Vermenigvuldig links en rechts met x_{t-k} en neem de verwachtingswaarde. Omdat $E(x_{t-k} a_t) = 0$ voor $k > 0$, volgt:

$$\rho(k) = \alpha_1 \rho(k-1) + \alpha_2 \rho(k-2) + \dots + \alpha_M \rho(k-M) \quad k > 0 \quad (9)$$

Als we $k = 1, 2, \dots, M+1$ in (9) substitueren krijgen we een stelsel vergelijkingen, die bekend zijn als de Yule-Walker vergelijkingen. (Yule 1927; Walker 1931).

$$\begin{aligned} \rho(1) - \alpha_1 \rho(0) - \dots - \alpha_M \rho(M-1) &= 0 \\ \rho(2) - \alpha_1 \rho(1) - \dots - \alpha_M \rho(M-2) &= 0 \\ \rho(M+1) - \alpha_1 \rho(M) - \dots - \alpha_M \rho(1) &= 0 \end{aligned} \quad (10)$$

Veronderstel nu dat de eerste $M+1$ waarden $\rho(0), \rho(1), \dots, \rho(M)$ van de autocovariantiefunctie bekend zijn. Substitutie van deze waarden in de eerste M vergelijkingen van (10) geeft de coëfficiënten $\alpha_1, \alpha_2, \dots, \alpha_M$.

De onbekende coëfficiënt $\rho(M+1)$ wordt nu bepaald door op te lossen:

$$\det \begin{bmatrix} \rho(1) & \rho(0) & \dots & \rho(M-1) \\ \rho(2) & \rho(1) & \dots & \rho(M-2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(M+1) & \rho(M) & \dots & \rho(1) \end{bmatrix} = 0 \quad (11)$$

Een vergelijking van (11) en (7) geeft exact dezelfde waarde voor $\rho(M + 1)$ als bepaald door het maximaliseren van de entropie van het proces.

Deze equivalentie betreft al de geëxtrapoleerde autocovariantie coëfficiënten. Geconcludeerd mag dus worden dat de MEM spectraal analyse equivalent is aan de aanpassing van een AR model aan het random proces. (Van den Bos, 1971). Het spectrum van het AR proces voor $\Delta t = 1$ kan worden verkregen door de Z getransformeerde van (8) te nemen.

$$X(Z) - X(Z) (\alpha_1 Z + \alpha_2 Z^2 + \dots + \alpha_M Z^M) = A(Z) .$$

$$|X(Z)|^2 = \frac{|A(Z)|^2}{|1 - \alpha_1 Z - \alpha_2 Z^2 - \dots - \alpha_M Z^M|^2} . \quad (12)$$

Substitutie van $Z = \exp(-i\pi f)$ in (12) levert het halve energiespectrum en dus is het energiespectrum van het lineaire proces:

$$S(f) = \frac{2\sigma_a^2}{\left[1 - \sum_{j=1}^M \alpha_j \exp(-i2\pi f j)\right]^2} . \quad (13)$$

Wat tevergelijken is met (3).

Het is interessant dat de representatie van een stochastisch proces door een AR model de representatie is die een maximum aan entropie weergeeft. Het is duidelijk uit de vorm (8) dat een andere interpretatie van een AR proces er één is die x_t identificeert met een waarde die voorspeld wordt uit de $M - 1$ voorgaande waarden van het proces met een voorspellingsfout gelijk aan a_t . Dus $\alpha_1, \alpha_2, \dots, \alpha_M$ is een M punts voorspellingsfilter (Whittle, 1963; Peacock and Treitel, 1969).

Herschrijf (8) als:

$$a_t = x_t - \alpha_1 x_{t-1} - \alpha_2 x_{t-2} - \dots - \alpha_M x_{t-M} .$$

Identificeer $1, -\alpha_1, -\alpha_2, \dots, -\alpha_M$ met een $M + 1$ punt voorspellingsfout-filter. De coëfficiënten $-\alpha_1, -\alpha_2, \dots, -\alpha_M$ worden gewoonlijk geschreven als $\gamma_1, \gamma_2, \dots, \gamma_M$ en het fout voorspellingsfilter wordt

1, 1, 2, ... M.

Het schatten van de AR coëfficiënten.

Om het MEM spectrum te berekenen met (13) of (3) moeten we eerste de lengte van het filter bepalen (M). Dit is gelijk aan de orde van het AR proces. Ten tweede moeten de coëfficiënten α_j zelf bepaald worden. Omdat de methode volgens Akaike voor de bepaling van M kennis over de α_j verondersteld, wordt de bepaling van de parameters eerst besproken. Er worden twee schatters gebruikt.

1. De Yule Walker schatters.

Een gewone methode om schatters $\hat{\alpha}_j$ van α_j te bepalen is door de Yule Walker vergelijkingen (10) op te lossen met als schatter $\hat{\rho}(k)$ voor de autocovarianties $\rho(k)$.

$$\hat{\rho}(k) = \frac{1}{N} \sum_{t=1}^{N-|k|} (x_{t-k} - m) (x_t - m) . \tag{14}$$

met $m = \frac{1}{N} \sum_{t=1}^N x_t .$

Deze methode geeft een fout die over het algemeen kleiner is dan al de andere schatters (Jenkins en Watts, 1968).

De UW vergelijkingen kunnen nu als volgt worden geschreven

$$\begin{bmatrix} \hat{\rho}(0) & \hat{\rho}(1) & \dots & \hat{\rho}(M-1) \\ \hat{\rho}(1) & \hat{\rho}(0) & & \hat{\rho}(M-2) \\ \vdots & \vdots & & \vdots \\ \hat{\rho}(M-1) & \hat{\rho}(M-2) & \dots & \hat{\rho}(0) \end{bmatrix} \cdot \begin{bmatrix} \hat{\alpha}_{M1} \\ \hat{\alpha}_{M2} \\ \vdots \\ \hat{\alpha}_{MM} \end{bmatrix} = \begin{bmatrix} \hat{\rho}(1) \\ \hat{\rho}(2) \\ \vdots \\ \hat{\rho}(M) \end{bmatrix} \tag{15}$$

met $\hat{\alpha}_{Mj}$ de j-de coëfficiënt van het M-de orde AR proces.

Een recursieve oplossing voor (15) kan worden verkregen met de Levinson (1947) en Durbin (1960) procedure. Hiervoor is het nodig een alternatieve vorm van de YW vergelijkingen op te stellen.

Vermenigvuldig (8) met x_t en neem de verwachtingswaarde.

Dit geeft met $E(x_t \cdot a_t) = E(a_t^2) = \sigma_a^2$,

$$\rho(0) = \alpha_1 \rho(1) + \alpha_2 \rho(2) + \dots + \alpha_M \rho(M) + \sigma_a^2 . \tag{16}$$

zodat nu de YW vergelijkingen (10) geschreven kunnen worden met (16).

$$\begin{bmatrix} \rho(0) & \rho(1) & \dots & \rho(M) \\ \rho(1) & \rho(0) & \dots & \rho(M-1) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(M) & \rho(M-1) & \dots & \rho(0) \end{bmatrix} \begin{bmatrix} 1 \\ -\hat{\alpha}_1 \\ \vdots \\ -\hat{\alpha}_M \end{bmatrix} = \begin{bmatrix} \sigma_a^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} P_{M+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (17)$$

P_{M+1} is de energie van het $M+1$ -punts fout-voorspellings AR filter γ_j .
 Bekeken wordt als voorbeeld de iteratie procedure voor $M=3$.
 Als $\hat{C}(3)$ de schatter is voor de covariantiematrix levert het linkerlid van (17), met gebruik maken van (15) voor $M=3$ en $M=2$:

$$\hat{C}(3) = \begin{bmatrix} 1 \\ -\hat{\alpha}_{31} \\ -\hat{\alpha}_{32} \\ -\hat{\alpha}_{33} \end{bmatrix} = \hat{C}(3) \begin{bmatrix} 1 \\ -\hat{\alpha}_{21} \\ -\hat{\alpha}_{22} \\ 0 \end{bmatrix} - \hat{\alpha}_{33} \begin{bmatrix} 0 \\ -\hat{\alpha}_{22} \\ -\hat{\alpha}_{21} \\ 1 \end{bmatrix} \quad (18)$$

En het rechterlid =

$$\begin{bmatrix} P_4 \\ 0 \\ \cdot \\ 0 \end{bmatrix} = \begin{bmatrix} P_3 \\ 0 \\ \cdot \\ 0 \end{bmatrix} - \hat{\alpha}_{33} \begin{bmatrix} \Delta_3 \\ 0 \\ \cdot \\ P_3 \end{bmatrix} \quad (19)$$

De iteratie wordt nu bepaald door de volgende betrekkingen die de Δ_3 en P_3 bepalen:

$$\begin{aligned} \Delta_3 &= \hat{\rho}(3) - \hat{\alpha}_{21}\hat{\rho}(2) - \hat{\alpha}_{22}\hat{\rho}(1) , \\ P_3 &= \hat{\rho}(0) - \hat{\alpha}_{21}\hat{\rho}(1) - \hat{\alpha}_{22}\hat{\rho}(2) .. \end{aligned}$$

Uit (19) volgt dan $\hat{\alpha}_{33} = \frac{\Delta_3}{P_3}$ en de recursie is compleet.
 Deze iteratie is gevoelig voor afrondingsfouten vooral als het AR proces bijna instationair is. Een bezwaar tegen de methode met YW schatter voor $\rho(t)$ is dat deze schatter (14) uitgaat van de veronderstelling dat $x_t = M$

voor $|t| > N$. Dit is in tegenspraak met het principe van maximale entropie.

2. De Burg-schatter.

Burg suggereerde een methode om de AR parameters te schatten, die geen eerste schatter van de autocovariantiefunctie vereist. De Burgmethode gebruikt een recursie die erg lijkt op YW schatter. Burg minimaliseert de totale fout in de M^{de} iteratie $\sum_{t=M}^N a_t^2$ naar de $\hat{\alpha}_{MM}$. Dus $\hat{\alpha}_{MM}$ wordt verkregen uit de volgende uitdrukking:

$$\frac{\partial S(\alpha_{MM})}{\partial \alpha_{MM}} = 0 \quad \text{met } S(\alpha_{MM}) = \sum_{t=M+1}^N (x_t - \hat{\alpha}_{M1}x_{t-1} - \hat{\alpha}_{M2}x_{t-2} - \hat{\alpha}_{MM}x_{t-M})^2 \quad (20)$$

Burg stelde voor om niet (20) te gebruiken daar de fout slechts bepaald wordt door de data met positieve tijdas te doorlopen, maar om de data in twee richtingen te doorlopen met een foutvoorspellingsfilter en deze totale foutenergie P_M te minimaliseren naar α_{MM} . Bijvoorbeeld voor een derde orde AR proces.

$$P_4 = \frac{1}{2(N-3)} \sum_{t=4}^N (x_t - \hat{\alpha}_{31}x_{t-1} - \hat{\alpha}_{32}x_{t-2} - \hat{\alpha}_{33}x_{t-3})^2 + (x_{t-3} - \hat{\alpha}_{31}x_{t-2} - \hat{\alpha}_{32}x_{t-1} - \hat{\alpha}_{33}x_t)^2 .$$

α_{33} wordt nu bepaald uit $\frac{\partial P_4}{\partial \hat{\alpha}_{33}} = 0$.

Belangrijk is dat geen veronderstellingen over de onbekende gegevens worden meegenomen. Deze schatter is dus consistent met het principe van maximale entropie.

Als $\hat{\rho}(0)$ op de gewone manier wordt berekend:

$$\hat{\rho}(0) = \frac{1}{N} \sum_{t=1}^N x_t^2 .$$

Volgen de resterende autocovariantieschatters recursief.

Uit (17) volgt: $\hat{\rho}(\tau) = \sum_{k=1}^{\dagger} \hat{\rho}(\tau - k) \hat{\alpha}_{\tau k}$; $\tau = 1, 2, \dots, M$. (25)

Andersen (1974) geeft een recursie schema.

De bepaling van de orde van het AR proces volgens Akaike (1970).
Gedefinieerd wordt een fout, de "Final prediction error" afgekort tot FPE, volgens:

$$FPE_M = E((x_t - \hat{x}_t)^2) .$$

Hierin is \hat{x}_t de schatter van x_t in iteratie-stap M.
Akaike bewees dat voor een M de orde proces:

$$FPE_M = (1 + \frac{M + 1}{N})\sigma^2 ,$$

met σ^2 de variantie van de innovatie van het proces.
Een schatter voor de innovatievariantie is volgens Akaike $\hat{\sigma}^2$.

$$\hat{\sigma}^2 = (\frac{N}{N - (M + 1)}) S_M^2 \text{ met } S_M \text{ volgens (20).}$$

Dit levert voor FPE_M

$$FPE_M = \frac{N + (M + 1)}{N - (M + 1)} S_M^2 . \tag{22}$$

Bij iedere stap wordt deze residu-som van kwadraten S_M^2 berekend en FPE_M bepaald. Als FPE_M minimaal is, als functie van M, geeft deze M de optimale AR schatter voor het proces x_t .

Als het gemiddelde niet vooraf verwijderd is geldt:

$$FPE_M = (\frac{N + M}{N - M}) S_M^2 .$$

De orde van het AR proces hangt af van het proces x_t dat we willen weergeven in het spectrale domein. Elke complexe component van Γ of $H(j\omega)$ heeft twee reële waarden nodig voor zijn definitie. Dus een M term filter, kan slechts $M/2 + 1$ discrete frequentie componenten beschrijven (Jensen en Ulrich, 1973). Dit is een conditie, die belangrijk is bij de analyse van tijdreeksen met de MEM, lso van het Akaike criterium.

Enkele verdere eigenschappen (Ulrich en Bisschop, 1975)

1. De schatter voor het spectrum is asymptotisch zuiver of $E(\hat{P}(f)) = P(f)$ als $\hat{P}(f)$ een schatter is voor $P(f)$.
2. De schatter is asymptotisch normaal.
3. $\text{Var.}(\hat{P}(f)) = 2\frac{M}{N} P^2(f)$, voor grote M en N , dus als het spectrum **glad** is. Deze variantie is dus vergelijkbaar met het ruwe periodogram.

Lacoss (1971) voerde numerieke berekeningen uit op de autocorrelatie-functie van een sinus met energie α en frequentie λ , met hierop ruis gesuperponeerd met energie 1. Dit signaal wordt bemonsterd met interval Δ en levert dan een correlatie-functie $C_n = \delta_n + \alpha \cos \lambda_n$. Hij berekende de spectraal schatter van ρ_n met o.a. de MEM en ook met de ruwe periodogramschatter.

Dit voor de discrete waarden $n = 0, \dots, N$ met $N \gg 1$ en $\alpha N \gg 1$. Dit resulteert in de navolgende tabel.

Schatter	Bandbreedte $W(H_z)$	Piekhogte P	W.P. (oppervlakte)
Ruwe P	$\frac{\sqrt{6}}{\pi \Delta N}$	$\frac{\alpha}{2}$	$\frac{\sqrt{6\alpha}}{2\pi \Delta N}$
MEM P_E	$\frac{4}{\pi \Delta N^2 \alpha}$	$\frac{\alpha^2 N^2 \Delta}{4}$	$\frac{\alpha}{\pi}$

De MEM heeft dus een groot oplossend vermogen, evenredig met N^2 .

De hoogte van de piek wordt niet zuiver weergegeven, wel het oppervlak onder de piek W.P. Het hoge oplossend vermogen van de MEM wordt gebruikt om periodiciteiten aan te tonen.

Om foutloze piekhoogten te krijgen wordt het ruwe MEM spectrum geïntegreerd. Dit wordt o.a. door Jensen en Ulrich (1973) gebruikt in de analyse van fluktuaties in Barnard's ster.

De MEM voldoet zeer goed. Ook bij korte reeksen. Bij frequenties kleiner dan $4/T$, moet evenals bij het periodogram rekening worden gehouden met verschuivingen naar de lage frequenties (Ulrich, 1972).

Chen en Stegen (1973) toonden aan dat een hogere orde filter M moet worden gebruikt als het ruisniveau toeneemt. Er bestaat een niveau van ruis, waarbij de MEM niet meer voldoet. Het aantal samples moet dan vergroot worden. Een computerprogramma in FORTRAN voor de AR parameters is te vinden in Ulrich en Bisschop (1975).

Gersch en Sharpe (1973) bekijken de FPE van een bekend oneindig autoregressief data schema. In het verleden is slechts de bruikbaarheid van FPE op een eindig AR schema getest.

Opmerkingen:

1. Een te klein AR filter geeft een te gladde schatter, die voorbijgaat aan het hoge oplossende vermogen van de MEM, terwijl een buitengewoon grote lengte, valse pieken doet ontstaan in het spectrum.
2. Naast de MEM bestaat er ook een Maximum Likelihood schatter voor het spectrum, afgekort MLM. Burg (1972) toonde aan dat MLM schatter uit de MEM af te leiden is en bovendien inferieur is in eigenschappen, zoals het oplossend vermogen.

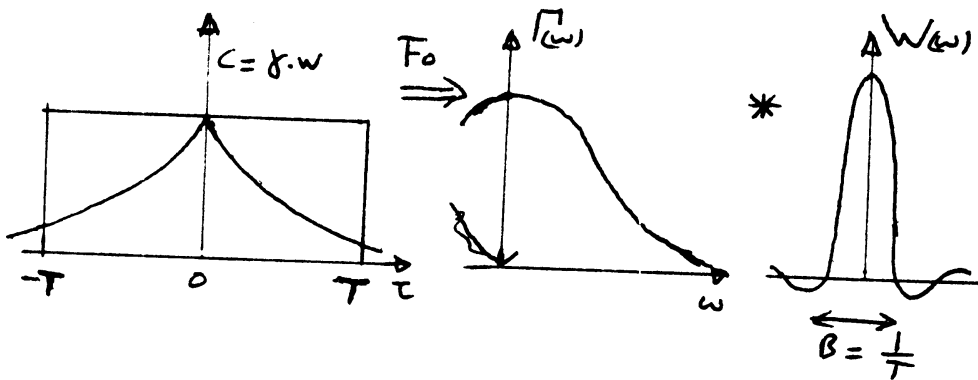
13. De exacte spectraalrepresentatie.

Deze techniek is in staat om exact het spectrum te berekenen van een (4de orde) stationair en "bandlimited" signaal.

Met het spectrum wordt de energieverdeling, behorende bij de gegeven meetlengte van de stochastische variabele, bedoeld, dus een monster van het echte spectrum.

De methode werd voor het eerst naar voren gebracht door Zagalsky (1967), die de methode baseerde op de oplossing van Barnes (1966) voor de reconstructie van een optische diffractie. De wiskundige achtergrond werd door Slepian, Pollak en Landau (1961) ontwikkeld. De methode berust op het feit, dat er spheroïde golffuncties bestaan, die orthogonaal zijn op een eindig, zowel als een oneindig tijdsinterval. Met deze functies is dan terugtransformatie uit het door tijdvenster T bekeken spectrum mogelijk.

Uitgegaan wordt van de bekende autocorrelatiefunctie, die afgebroken is, zeg met tijd T. Er wordt dus een rechthoekig tijdvenster over de a.c.v.f. gelegd, waarmee deze vermenigvuldigd wordt. In het spectrale domein betekent dit convolutie met een spectraalvenster.



Figuur 21. Het spectrum $\Gamma(\omega)$ wordt door een venster bekeken, waardoor de bandbreedte van het venster het oplossend vermogen bepaalt, maar ook ontstaan er ernstige fouten door de eindige Fouriertransformatie, vooral bij lage frequenties.

Door het venster gezien wordt het spectrum $C(\omega)$, wat biased is.

$$C(\omega) = \int_{-\infty}^{+\infty} \frac{\sin T(\omega - x)}{\pi(\omega - x)} \Gamma(x) dx . \quad (13.1)$$

De exacte spectraalrepresentatie behelst nu een terugtransformatie, met behulp van het ontbinden van $\Gamma(\omega)$ in spheroïde golffuncties Ψ_r . Als W de bandbreedte van Γ is leidt dit volgens de benadering van Zagalsky (1967) tot de oneindige som:

$$\Gamma(\omega) = \sum_{n=0}^{\infty} \int_{-W}^{+W} \frac{\Psi_n(\omega) \cdot \Psi_n(x)}{\lambda_n^2} C(x) dx . \quad (13.2)$$

De Ψ_i en λ_i zijn de eigenfuncties, geheten spheroïde golffuncties, en eigenwaarden van de integraalvergelijking

$$\lambda_i \cdot \Psi_i(\omega) = \int_{-W}^{+W} \frac{\sin T(\omega - x)}{\omega - x} \Psi_i(x) dx . \quad (13.3)$$

Zowel Ψ_i als λ_i zijn afhankelijk van het product $W \cdot T$. (Slepian en Pollak). Uit 13.2 blijkt, dat $\Gamma(\omega)$ te benaderen is met willekeurige kleine fout, door de eindige som te berekenen over de integraal, uit-

gaande van het ruwe periodogram $C(\omega)$.

De spheroïde golf functies hebben de eigenschap, dat ze orthogonaal zijn op het oneindige frequentie-interval en op de eindige bandbreedte.

$$\int_{-\infty}^{+\infty} \Psi_i(\omega) \cdot \Psi_j(\omega) d\omega = \begin{cases} 1 & \text{als } i=j \\ 0 & \text{als } i \neq j \end{cases} \quad i, j = 0, 1, 2, \dots \quad (13.4)$$

$$\int_{-w}^{+w} \Psi_i(\omega) \cdot \Psi_j(\omega) d\omega = \begin{cases} \lambda_i & \text{als } i=j \\ 0 & \text{als } i \neq j \end{cases} \quad i, j = 0, 1, 2, \dots \quad (13.5)$$

Uitgaande van Slepian en Pollak (1961) is het mijns inziens ook mogelijk en zelfs éénvoudiger, direct de afgebroken a.c.v.f. $\gamma(\tau)$ te ontbinden in spheroïde golf functies en vervolgens de oneindige Fouriertransformatie over deze golf functies te bepalen. Dit levert dan het spectrum met willekeurig kleine fout.

Ontbindt $f(t) = \gamma(\tau)$ in spheroïde golf functies $\Psi_n(t)$.

$f(t)$ is bekend over het interval $-T < t < T$.

λ_i en $\Psi_i(t)$ zijn nu eigenwaarden en eigenfuncties van de vgl. 13.2.

De functies $\Psi_i(t)$ hebben dubbele orthogonaliteit:

$$\int_{-\infty}^{+\infty} \Psi_i(t) \cdot \Psi_j(t) dt = \begin{cases} 1 & \text{als } i=j \\ 0 & \text{als } i \neq j \end{cases} \quad i, j = 0, 1, 2, \dots$$

$$\int_{-T}^{+T} \Psi_i(t) \cdot \Psi_j(t) dt = \begin{cases} \lambda_i & \text{als } i=j \\ 0 & \text{als } i \neq j \end{cases} \quad i, j = 0, 1, 2, \dots$$

Ontbindt $f(t)$: $f(t) = \sum_0^{\infty} \alpha_n \Psi_n(t)$,

$$\text{met } \alpha_n = \int_{-\infty}^{+\infty} f(t) \Psi_n(t) dt = \frac{1}{\lambda_n} \int_{-T}^{+T} f(t) \Psi_n(t) dt .$$

De coëfficiënten α_n voor het oneindige interval kunnen dus bepaald worden uit de waarden van $f(t)$ in het eindige interval.

De oneindige som kan benaderd worden door de eindige som, daar geldt:

$$\lim_{N \rightarrow \infty} \int_{-\infty}^{+\infty} (f(t) - \sum_0^N \alpha_n \psi_n(t))^2 dt = 0 \quad (\text{Landau en Pollak, 1961}).$$

De fout, die gemaakt wordt, als $f(t)$ benaderd wordt door

$$f_N(t) = \sum_0^N \alpha_n \psi_n(t) \text{ gelijk is aan}$$

$$\int_{-\infty}^{+\infty} (f(t) - f_N(t))^2 dt = \sum_{n+1}^{\infty} \alpha_n^2.$$

Deze kan zo klein gemaakt worden als gewenst, door N voldoende groot te nemen. De exacte spectraalrepresentatie vermijdt de moeilijkheden betreffende het venster. Vooral belangrijk is de zuiverheid bij lage frequenties, dit tot nul toe. Het verkregen spectrum is volkomen unbiased en vertoont geen frequentieverschuivingen.

Het oplossend vermogen kan in principe oneindig groot gemaakt worden en kent geen begrenzing door een spectraalvenster.

Er zijn nog geen computerprogramma's aanwezig, naar mijn weten.

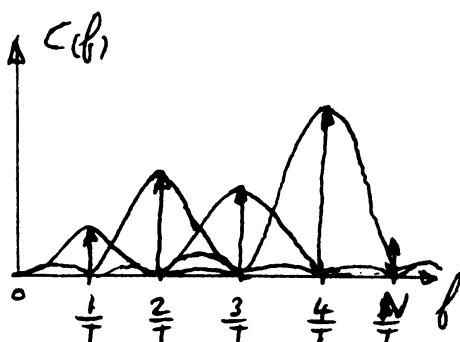
14. Het bepalen van een spectrum door filteren.

Bij discrete Fouriertransformatie wordt $C(f)$ in een $\frac{N}{2}$ aantal punten bepaald

$$C(f) = \frac{1}{T} \left[\int_{-\frac{T}{2}}^{+\frac{T}{2}} x(t) \cdot e^{-j2\pi ft} dt \right]^2.$$

$C(f)$ heet de ruwe periodogram schatter en heeft een grote variantie.

$\text{Var.}\{C(f)\} = \Gamma^2(f)$, omdat $C(f)$ opgebouwd is uit onafhankelijke discrete staken van schatterwaarden.



Figuur 22.

De ruwe periodogramschatter met N onafhankelijk staken.

Het spectraalvenster $W(f-g)$ is nul waar andere staken staan.

De staken vertegenwoordigen over eindige meettijd T , het filteren over een frequentie gebied van de breedte $\frac{2}{T}$ van het Bartlett spectraalvenster.

$$W_B(f) = T \left(\frac{\sin \pi fT}{\pi fT} \right)^2 .$$

De staken zijn dus onafhankelijk als de afstand tussen de discrete frequenties, $\Delta f \gg \frac{2}{T}$.

De variantieverhouding is dan gelijk één.

Als het ruwe periodogram gesmootht wordt, bijvoorbeeld door een venster toe te passen wordt een afhankelijkheid geïntroduceerd, met als gevolg een biased spectrum, met echter een kleinere variantie. Het filter W_B wordt als het ware verbreed, terwijl de Δf dezelfde blijft.

Al deze technieken van vensters en filter zijn te vertalen in het analoge en digitale filteren. Uit het voorgaande blijkt dat het geen zin heeft een analogo of digitaal filter toe te passen op een tijdreeks van lengte T met een bandbreedte van $2/T$ of kleiner.

Minimaal moet de bandbreedte van het filter gelijk aan die van W_B worden gekozen ($\frac{2}{T}$).

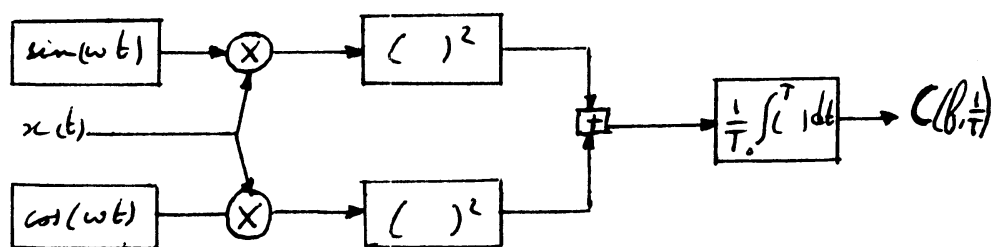
Dit komt analoog, digitaal of wiskundig neer op het vermenigvuldigen van $x(t)$ met een sinus en cosinus voor $t = \{0 \dots T\}$.

Kwadratering levert dan de energie in bandbreedte $2/T$.

Dit is dus exact hetzelfde als directe Periodogram bepaling.

$$C(f, \frac{1}{T}) = \frac{1}{T} \left[\int_{-\frac{T}{2}}^{+\frac{T}{2}} x(t) \cdot e^{-j2 \pi ft} dt \right]^2 .$$

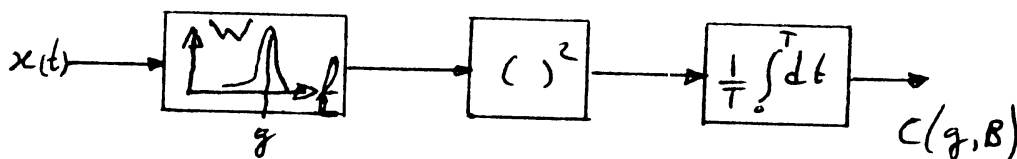
Dit scherp bandfilteren kan analoog eenvoudig worden opgezet.



Figuur 23. Het analoog bandfilteren met bandbreedte $\frac{2}{T}$.

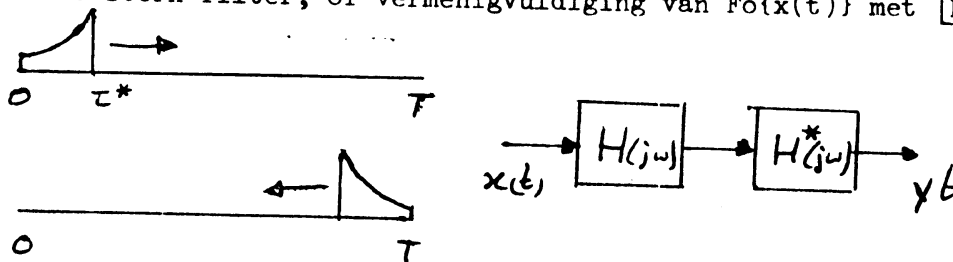
Kommentaar bij figuur 20. Nodig is een sinusgenerator, twee vermenigvuldigers \otimes , twee kwadraten $(C)^2$ en een opteller \oplus . Bovendien een integrator (of een lowpassfilter met hoge afkapproefrequentie).

Om "smoothing" te verkrijgen met B kan een specifiek filter groter dan $\frac{2}{T}$ gekozen worden. De variantie wordt dan t.o.v. het ruwe periodogram met een factor $B \cdot \frac{T}{2}$ verkleind.



Figuur 24. Het analoge passbandfilter met variabel filter om een gladgestreken periodogram te verkrijgen.

Hetzelfde kan ook digitaal worden uitgevoerd op een discrete reeks van getallen $x(n)$. Als simpel met een sinus en cosinus vermenigvuldigd wordt verkregen we een ruw periodogram. Dit is niets anders dan discrete Fouriertransformatie. Een gladder spectrum kunnen we verkrijgen door een digitaal filter op de reeks los te laten, met bandbreedte $B > \frac{2}{T}$. Evenals een analog filter bezit het digitale filter een insteltijd τ^* , waarna het gefilterde, foutloze waarden aflevert. Deze insteltijd is te vinden uit de geheugen- of pulsresponsie $k(\tau)$ van het filter. Bij eindige discrete reeks gaat dit deel τ^* verloren voor de bepaling van het spectrum. Een mogelijkheid om dit te omzeilen is om de discrete reeks zowel met positieve als negatieve tijden te doorlopen. Optimaal wordt van alle informatie gebruik gemaakt als $\tau^* < T/2$, zie figuur 25. Sterk aangeraden wordt dus om de pulsresponsie $k(\tau)$ te berekenen als Fourierteruggetransformeerde van de complexe overbrengingsverhouding $H(j\omega)$. Doorlopen in twee richtingen betekent filteren met een tweemaal zo sterk filter, of vermenigvuldiging van $Fo\{x(t)\}$ met $[H(j\omega)]^2$.



Figuur 25. Het tweezijdig doorlopen van de discrete reeks door een digitaal filter maakt optimaal gebruik van alle waarden als $\tau^* < T/s$.

Deze τ^* wordt bepaald door de steilheid van het toe te passen filter. Hoe steiler het filter hoe groter de insteltijd τ^* .

Om steiler te filteren kan de reeks meerdere malen worden doorlopen.

Als $x(t)$ tweemaal of een even aantal malen negatief en positief wordt doorlopen vertoont de gefilterde reeks geen faseverschuiving in tegenstelling tot de éézijdig doorlopen reeks.

Opmerkingen: De theorie en het ontwerp van digitale filters wordt hier niet behandeld. Verwezen wordt naar een werk van twee pioniers op dit gebied nl. Radar en Gold (1965). Met behulp van digitaal high-pass filteren kunnen lage frequenties, die aanleiding kunnen geven tot harmonischen bij eindige Fouriertransformatie, vooraf weggefilterd worden. Wel moet rekening worden gehouden met de insteltijd. Qua eigenschappen is het analoog en digitaal filteren met een frequentie onafhankelijk filter geheel dezelfde als de bepaling van een spectrum met behulp van de venstermethode.

15. Vergelijking van de verschillende methoden.

Als vergelijkende maatstaven worden gekozen:

1. De Bias: $b = E(\bar{C}(f)) - \Gamma(f)$.

Als de bias naar nul gaat, hebben we te maken met een asymptotisch-zuivere schatter. Dit wordt aangegeven door een slangetje \sim .

2. De variantieverhouding: $\text{Var.}(\bar{C}(f))/\Gamma^2(f)$.

3. De bandbreedte van het filter B , wat het oplossend vermogen bepaalt. Bij de ruwe periodogramschatter hebben we te maken met een Bartlett venster. Bij de venstermethode kiezen we een Tukey of Parzenvenster met lengte M .

4. De laagste frequentie die kan worden gedetecteerd, zonder verschuiving naar lage frequenties met een maximale verschuiving van 5%. *

5. De benodigde hoeveelheid bewerkingen.

Een bewerking is één optelling en één vermenigvuldiging.

Steeds wordt het snelste algoritme gebruikt, b.v. FFT.

* Voor de MEM: zie Aangangsel A1.

16. Tabel van vergelijkende maatstaven.

	bias	variantie ratio	bandbreedte (Hz)	laagste freq. (Hz)	aantal bewerkingen
ruwe periodogram	0	1	$\frac{2}{T}$	$\frac{4}{T}$	$2N^2 \ln N$
gefilterd periodogram	$\frac{C}{M^2} \Gamma_{xx}(2)$	$\frac{M}{N}$	$\frac{N}{M} \cdot \frac{2}{T}$	$\frac{N}{M} \cdot \frac{4}{T}$	$2N^2 \ln N$
Jenkinson periodogram	0	≈ 1	$\lambda \cdot f_m$	$\frac{4}{T}$	$2N^2$
VRL	0	≈ 1	$< \frac{2}{T}$	$\frac{4}{T}$	$\frac{4}{3} N^2 \ln N$
MLM	0	$\frac{M}{N}(\text{glad})$	$\frac{1}{\Delta N} \frac{1}{\sqrt{\alpha N}}$	$\frac{4}{T}$	N^2
MEM	0	$\frac{2M}{N}(\text{glad})$	$\frac{1}{\Delta N^2 \alpha}$	$\frac{4}{T}$	$N(N+1)$
exacte spectraal representatie	0	?	0	0	
(digitaal) filteren				$\frac{4}{T}$	

Opmerking: N wordt gebruikt om de discrete reekslengte weer te geven.

$T(\text{sec})$ is de eindige meettijd van de reeks.

M is de discrete lengte van het venster of het discrete AR filter bij de AR methoden.

Δ is het bemonsteringsinterval in seconden.

Bij de MLM en MEM wordt de bandbreedte van een sinus met amplitude α bekeken (Lacoss, 1971).

De bias is bij de venstermethode afhankelijk van de tweede afgeleide van $\Gamma(f) = \Gamma^{(2)}(f) = 0$ als Γ is glad.

De invloed van de vorm van het venster wordt weggelaten, slechts de orde van ~~g~~ grootte wordt aangegeven.

17. Toepassingsgebieden en conclusies.

De vraag welke de beste methode is, is niet zonder meer te beantwoorden. De bruikbaarheid van een methode hangt zowel af van het beschikbaar zijn van de hoeveelheid informatie (lengte van de tijdreeks), als van de statistische eigenschappen, die het spectrum worden toegewenst. Onderscheiden worden twee gevallen:

1. Men kan beschikken over een vrijwel oneindige tijdreekslengte T.

In principe is het spectrum nu exact te bepalen, met bias, variantie, bandbreedte zo klein als men wenst. Bovendien kunnen zeer lage frequenties gedetecteerd worden. Natuurlijk is voorop gesteld dat de reeks stationair is in het meetinterval.

De gevolgde methode is nu de gevensterde periodogram-methode, die vrijwel overal beschikbaar is in de vorm van een computerroutine.

Men kan ook andere, b.v. de MEM gebruiken, hoewel deze minder voor de hand liggend zijn, want de a.c.v.f. is over een voldoende groot gebied nauwkeurig bekend om het spectrum te berekenen door Fouriertransformatie. Over het algemeen kosten deze alternatieven meer rekentijd bij gelijke kwaliteit.

2. Men beschikt slechts over een eindige, beperkte tijdreekslengte T.

Afhankelijk van de gevolgde methode, zijn beperkingen opgelegd aan bias, variantie, bandbreedte en laagst te detecteren frequentie.

De beperkingen zijn bovendien nog afhankelijk van de vorm van het spectrum. Pieken vereisen een minimale bandbreedte en als gevolg daarvan b.v. bij de venstermethode een bias (afhankelijk van de 2e afgeleide van $\Gamma(f)$ en een grotere variantie (variantie \times bandbreedte = konstant).

Als pieken met kleinere breedte dan $\frac{2}{T}$ gedetecteerd moeten worden, moet overgestapt worden op de exacte spectraalrepresentatie, die het frequentie venster elimineert.

Aanbevelingen:

De bruikbaarheid van de methoden hangt vooral af van de laagste te bepalen frequentie in verhouding tot de lengte T van de tijdreeks. Dit moet geschieden met zekere nauwkeurigheid in frequentie en met een variantie, die klein genoeg is.

Als norm of maximale fout in een frequentie verschuiving wordt 5% aangehouden. Dit correspondeert met een laagst te bepalen frequentie van $\frac{4}{T}$ (Toman, 1965). Voor redelijke smoothing van het periodogram is een middeling van 10 schatters vereist om tot een redelijk uitziende gevensterde schatter te komen.

Met de venstermethode leidt de norm van 5% verschuiving tot een redelijk spectrum bij frequenties hoger dan $\frac{40}{T}$ (Hz).

Als men geïnteresseerd is in lagere frequenties dan deze waarde, is de meest aangewezen method de maximale entropie methode, die tegenwoordig tot volwassenheid ontwikkeld is. De laagste frequentie, die hiermee, met de norm van 5% kan worden bepaald is $\frac{4}{T}$ (Hz).

Wil men uitgaande van de reeks T lagere frequenties zuiver bepalen, dan moet men de Exacte spectraal representatie ontwikkelen en toepassen.

Samengevat: aanbevolen wordt:

voor frequenties $f > \frac{40}{T}$, de venstermethode;

voor frequenties $f > \frac{4}{T}$, de Maximale entropie methode;

voor frequenties $f > 0$, de exacte spectraal representatie.

Voor het KNMI, waar slechts de venstermethode als computerroutine aanwezig is (1979), lijkt het me daarom zinvol om de MEM als computerroutine in te voeren. Een FORTRAN programma is te vinden in Ulrich en Bischof (1975). Deze MEM heeft een groter bereik, wat vooral belangrijk is voor de lage frequenties en kan zich zeker meten met de venstermethode in het bereik $f > \frac{40}{T}$, met een hoger oplossend vermogen en minder bias. Tevens moet overwogen worden om de Exacte spectraal representatie te ontwikkelen tot een routine; misschien kan deze eens nog concurrerend worden voor de MEM in het bereik $f > \frac{4}{T}$.

Onder deze frequentie heeft deze spectraal representatie zeker kwaliteiten, sterker nog, het is de enige bekende manier om van deze lage frequenties een unbiased spectrum te verkrijgen zonder frequentieverschuivingen.

Naast de MEM en de Exacte spectraal representatie is het misschien nuttig om voor lage frequenties het Jenkinsonperiodogram toe te passen (Jenkinson, 1977).

Zonder enige moeite kan ook de VRL techniek voor frequenties $f > \frac{4}{T}$ gebruikt worden. Deze is echter minder goed dan het Jenkinsonperiodogram en hij verbruikt veel meer rekentijd.

18. Referenties bij deel II.

- | | |
|----------------------------|--|
| Cooley en Tukey, 1965 | - An algorithm for the machine calculation of complex Fourier series, Math. of Comp. 1g, AM 1965. |
| Koopmans, 1974 | - The Spectral analysis of time series, Acad. press., New York, London, 1974. |
| Jenkins en Watts, 1968 | - Spectral Analysis and its applications, 1968 Holden Day, San Francisco. |
| Jenkinson, 1977 | - Met. Office B, Branch Memorandum no. 57. |
| Schickedanz en Bowen, 1975 | - Computation of Climatological Power Spectra using VRL. Preprint Fourth Conf. on Prob. and St. in Am. Sc., Nov. 1975. |

- Capon, 1970 - Probability distributions for estimators of freq. wave number spectrum, Proc. IEEE, VS 8, 1970.
- Ulrich en Bishop, 1975 - Maximum Entropie Spectral Analysis and autoregressive decomposition, Rev. Geoph. Spaceph., V 13, 1975.
- Lacoss, 1971 - Data adaptive spectral analysis methods, Geoph. V 36, No. 4, 1971.
- Yule, 1927 - On a method of inv. periodicities in disturbed series. Phil. Trans. Roy. Soc., London, Ser. A, 226, 267-298, 1927.
- Walker, 1931 - On periodicity in series of related terms, Proc. Roy. Soc., London, Ser. A, 131, 518-532, 1931.
- Jaynes, 1963 - New Engineering applications of information theory. Proc. First Symp. En. Appl. Rand. for Tr. Prob., 1963.
- Jaynes, 1968 - Prior Probabilities, IEEE, T.S.C.C., SEC-4, 227-241, 1968.
- Jansen en Ulrich, 1973 - An analysis of the perturbations on Barnard's star, Astr. J. V., 78, 1973.
- Akaike, 1970 - Statistical prediction identification. Ann. Inst. Statist. Math., V 22, 1970.
- Anderson, 1974 - On the calculation of filter coeff. for MEM, Geophysics, 39, 69-72, 1974.
- Fryer and all, 1975 -- Deconvolution and Spectral Estimation using F.P.E., Geoph., V 40, 1975.
- Chen and Stegen, 1974 - Experiments with Max. Entropie Power spectra of Sinoids, J. Geoph. Res., V 79, 1974.
- Gersch and Sharpe, 1973 - Estimation of Power spectra with finite order AR models, IEEE T.A.C., 1973.

- Zagalsky, 1967 - Exact Spectral Representation of Truncated Data, Proc. Letters of IEEE, Jan. 1967.
- Barnes, 1966 - Object restoration in a diffraction-limited imaging system, J. Opt. Soc. Am., V 56, 1966.
- Slepian and Pollak, 1961 - Prolate Spheroidal wave functions. Fourier analysis and uncertainty I. Bell. Syst. Tech. J., V 40, 1961.
- Landau and Pollak, 1961 - Prolate Spheroidal wave functions II. Bell Syst. Tech. J., V 40, 1961.
- Radar and Gold, 1965 - Digital Filter Design Techniques in the frequency domain, Proc. IEEE, V 55, 1965.
- Toman, 1965 - The spectral shifts of truncated sinusoids, J. Geoph. Res., V 70, 1965.
- Shannon, 1948 - A mathematical theory of communication, Bell. Syst. Tech. J. 27, 379-423, 1948.
- Smylie and all, 1973 - Analysis of irregularities in Earth's rotation, Methods of Computational Physics, Vol. 13, 391-430, New York.
- Van den Bos, 1971 - Alternative interpretation of MESA, IEEE Trans. Inf. Theory, I T-17, 1971.
- Whittle, 1963 - Prediction and Regulation by linear least-squares methods, D. van Noststrand, Princeton, N.J., 1963.
- Peacock and Treitel, 1969 - Predictive deconvolution, Theory and practice, Geoph. J., 33, 347-366, 1973.
- Levinson, 1947 - The Wiener RMS error criterium in filter design and prediction, J. Math. Phys., 25, 261-278, 1947
- Durbin, 1960 - The fitting of time series. Rev. Int. Inst. Statist., 28, 233-244, 1960.