



PROJETO DE FIM DE CURSO

**ESTUDO
DE SÉRIES TEMPORAIS
ASTRONÔMICAS DESIGUALMENTE
ESPAÇADAS:
MÉTODOS E APLICAÇÕES**

Aluno: Carlos Eduardo Lohse Rezende

*Orientadores: Jorge de Albuquerque Vieira
Encarnación Amelia Martínez Gonzalez*

Astronomia- UFRJ



ERRATA:

- No resumo, onde se lê "...classificações características.", leia-se "...classificações e características."
- Na página 6, última linha, onde se lê "...são coeficientes.", leia-se "...são coeficientes)."
- Na página 9, linha 16, onde se lê "...halteres...", leia-se "...haltere..."
- Na página 14, linha 17, onde se lê "Desprezar a componente horizontal previne uma certa tendência em regiões povoadas da amostra...", leia-se "Desprezar a componente horizontal evita uma certa tendência introduzida por regiões povoadas da amostra..."
- Na página 29, penúltima linha, onde se lê "...(iste é...)", leia-se "...(isto é...)".
- Na página 36, linhas 7 e 8 (fórmulas para A_G e B_G), na expressão $\text{sen } c(v\tau_n)$, leia-se $\text{sen } c(v'\tau_n)$

ÍNDICE:

	Página
1- Introdução	1
2- Busca de Periodicidade- Descrição de Métodos de Estudos de Séries Temporais para dados Desigualmente Espaçados	9
2.1- Sistemas Variáveis Com Periodicidade	9
2.2- Métodos Para Determinação de Períodos: Um Resumo	13
a) Métodos <i>String-Lenght</i>	13
b) Métodos <i>Phase-Binning</i>	15
c) Transformada de Fourier, Técnicas de Periodograma e Mínimos Quadrados	18
2.3- Complicações	20
a) <i>Aliasing</i>	20
b) Estimativa de Ruído	21
c) <i>Prewhitening</i> e Detecção de Múltiplos Períodos	22
3- Algoritmos Para Métodos Seleccionados	23
3.1- Método Jurkevich	24
a) Descrição do Método	24
b) O Programa	27
3.2 - Método Ponman (Mínimos Quadrados)	28
a) Descrição do Método	29
b) Aproximações e Erros	32

	Página
c) Programas	36
4- Aplicações e Discussão de Resultados	38
4.1- As Séries Temporais	38
4.2- Aplicação do Método Jurkevich às Séries Temporais	39
a) Dados Sintéticos de VW CEPHEI	39
b) Dados Sintéticos Senoidais	41
c) Dados Observados de HD 93044	41
4.3- Aplicação do Método Ponman às Séries Temporais	43
a) Dados Sintéticos de VW CEPHEI	43
b) Dados Sintéticos Senoidais	45
c) Dados Observados de HD 93044	46
4.4- Discussão dos Resultados	47
5- Conclusões	48
6- Agradecimentos	48
7- Bibliografia	49

*"Entre a mente que planeja
e as mãos que constroem,
tem que haver um mediador
e este é o coração."*

(Thea Von Harbou, *Metropolis*, 1926)

Dedicado à atriz alemã .
Brigitte Helm
que retratou com perfeição
as faces do Bem e do Mal.

RESUMO:

É apresentado, na introdução, um breve ensaio à respeito de séries temporais, suas classificações características. Posteriormente, aborda-se o problema de determinação de periodicidades em séries temporais astronômicas, desigualmente espaçadas. Uma breve descrição dos tipos de métodos de determinação de períodos é apresentada. Dois destes métodos são escolhidos e aplicados à três séries temporais, com o objetivo de se analisar qual método é mais ou menos eficiente para esta ou aquela série. É feita uma análise dos resultados obtidos e, por fim, as conclusões do trabalho obtidas a partir da análise dos resultados.

1- INTRODUÇÃO:

Em ciência, geralmente as teorias são formuladas apoiadas em evidências e testadas posteriormente em experimentos. Estes experimentos costumam ser reproduzidos em laboratórios e em condições totalmente controladas pelo cientista, o que proporciona testes conclusivos e teorias seguras, dando o respaldo de confiança de que a ciência necessita para continuar sendo uma instituição poderosa e eficiente. Entretanto, em algumas ramificações da ciência, estes procedimentos (experimentos efetuados para que se testem as teorias) não são possíveis de serem realizados. Tudo fica no campo da especulação e da elaboração de uma teoria que melhor explique o que estamos vendo sem nunca podermos provar experimentalmente os fatos. É o caso das ciências observacionais onde Astronomia está incluída. Não podemos, por exemplo, provar experimentalmente a existência do Big-Bang (se é que ele existe); talvez a teoria de ondas de densidade ou matéria escura expliquem a forma de certos sistemas galácticos, talvez não. O sucesso da teoria depende do fato dela ser convincente ou não. Nos é impossível ter o objeto de nosso estudo (no caso estrelas, galáxias, etc.) testados em laboratório. Esta, infelizmente, é uma limitação que a ciência observacional nos impõe e à qual precisamos aprender a conviver.

Outro problema existente em Astronomia e que complica a formulação de teorias convincentes é o tempo de vida que os objetos estudados possuem. A vida do cientista é muito menor que a do objeto e, portanto, certas variações ou propriedades deste podem não ser detectadas. Além disso, o pesquisador não pode estudar uma só estrela a vida inteira, por exemplo. Esta é uma complicação adicional que deve ser resolvida para se garantir teorias o mais adequadas possíveis a sistemas que só permitem o estudo via observação.

Imaginemos agora se dispuséssemos de uma ferramenta que, aplicada a um conjunto de observações num intervalo de tempo muito menor que o tempo de vida do objeto (uma estrela, por exemplo), nos desse uma previsão confiável de seu comportamento tanto no futuro quanto no passado. O problema da necessidade do excesso de observações estaria sanado e o cientista poderia observar outros sistemas, aumentando a sua produção científica. Uma previsão do comportamento do objeto daria ao teórico a chance de elaborar uma teoria mais segura sobre o seu comportamento. Esta ferramenta, para a alegria dos observadores e dos teóricos, existe e é a análise e modelagem de séries temporais. Uma série temporal é um conjunto de observações, sequenciais no tempo, de um fenômeno; entretanto, a variável tempo pode ser substituída por qualquer outra como espaço, profundidade, etc. As observações vizinhas são, em geral, dependentes e o estudo de uma série temporal consiste em analisar e modelar esta dependência. A série temporal pode também ser entendida como uma das realizações possíveis de um processo estocástico, isto é, dado um conjunto de observações X_i e seus respectivos tempos de observação t_i (onde $t_i \in T$, sendo T o conjunto de índices tempo no nosso caso), poderemos, num processo estocástico, ter vários valores de X_i para cada t_i (um processo estocástico pode ser conceituado como uma coleção de variáveis aleatórias ordenadas no tempo). Logo, para cada t_i temos que X_i é uma variável aleatória com uma função de distribuição. Isso gera um conjunto de várias séries temporais prováveis onde só uma ocorre. Este conjunto denomina-se *ENSEMBLE*. Deve-se ressaltar que numa série temporal, existe a dependência entre cada ponto X_i observado para diferentes t_i .

Os dois objetivos principais no estudo de séries temporais são:

- entender o mecanismo gerador da série temporal

- previsão do comportamento futuro do sistema

Para isso, os problemas de interesse são:

- a) descrever o comportamento da série via a construção de histogramas, gráficos, etc.
- b) investigar o mecanismo gerador da série temporal
- c) fazer previsões futuras a partir de valores passados
- d) busca de periodicidades

Podemos introduzir em nosso estudo o conceito de sistema dinâmico caracterizado por uma série de entrada $X(t)$, uma série de saída $Y(t)$ e uma função de transferência $v(t)$.

$$X(t) \longrightarrow v(t) \longrightarrow Y(t)$$

Um exemplo de relação entre saída e entrada são os sistemas lineares:

$$Y(t) = \sum_{u=0}^{\infty} v(u) X(t-u)$$

Aqui, os problemas de interesse são:

- a) estimar a função de transferência $v(t)$ conhecidas a série de entrada $X(t)$ e saída $Y(t)$
- b) fazer previsões da série $Y(t)$ com o conhecimento de $X(t)$ e $v(t)$
- c) estudar o comportamento do sistema, simulando-se a série de entrada
- d) controlar a série de saída $Y(t)$ de modo a trazê-la o mais próximo possível de um valor desejado ajustando-se convenientemente $X(t)$; este controle é necessário devido às perturbações que normalmente afetam um sistema.

As séries temporais podem ser classificadas a partir da natureza da observação do fenômeno X e do conjunto de índices T (tempo, espaço, etc.) como:

- discretas, onde T é um conjunto finito de pontos $T=1, 2, \dots, N$. Ex.: valor anual de exportações de 1950 a 1980.

- contínuas, onde T é um intervalo finito de pontos ($T[t: 0 < t < T]$). Ex.: Medições durante dois minutos de um eletrocardiograma.

- multivariada, onde temos vários conjuntos de observações X que possuem certa dependência entre eles: $[X_1(t), \dots, X_k(t)], t \in T$. Ex.: pressão uterina da mãe ($X_1(t)$) e batidas do coração do feto ($X_2(t)$) em trabalho de parto (pode ser discreta ou contínua).

↙ - multidimensional, onde \vec{r} é um vetor. Ex.: $[X(t, r, l), \vec{r} \in T]$ onde X é a altura de um ponto do oceano, t é o tempo, r é a altitude e l é a longitude.

As observações X podem então ser discretas ou contínuas, multivariadas ou univariadas e o índice tempo pode ser discreto ou contínuo, unidimensional ou multidimensional. Neste trabalho, utilizaremos séries univariadas discretas e unidimensionais.

O estudo clássico de séries temporais ainda decompõe a série em 4 componentes:

a) Componente sazonal: Muitas séries exibem uma variação periódica. Elas podem ser medidas e extraídas dos dados fornecendo uma série não-sazonal. Por exemplo, medidas de temperatura exibem uma variação anual periódica baixando durante o inverno e aumentando no verão.

b) Componente cíclica: São variações também periódicas mas seu período não é fixo como no caso da componente sazonal. Entretanto, este período pode ser previsto. Exemplo: ocorrência de surtos de sarampo que obedecem períodos de cerca de 2,5 a 3 anos.

c) Componente de tendência: São efeitos a longo prazo na média dos pontos da série. A dificuldade está em se definir longo prazo. Por exemplo, variações climáticas às vezes exibem

variações com períodos de 50 anos. Se tivermos 20 anos de observações, esta variação de 50 anos será confundida com tendência.

d) Componente irregular: Se removermos da série as 3 componentes acima, obtemos uma série de resíduos livre de qualquer variação regular. Esta série então é analisada à luz de um modelo probabilístico em que a noção de dependência é explorada.

Existem dois aspectos principais a serem considerados no estudo de séries temporais: a análise e a modelagem.

A principal razão para se modelar uma série temporal é tornar possível a previsão de valores futuros. Procura-se obter um modelo matemático que descreva o sistema fielmente. Este modelo deve ser de fácil manipulação matemática e altamente eficiente e funcional.

O objetivo da análise de uma série temporal é sumarizar as propriedades da série e caracterizar seu comportamento, identificando ou sugerindo um modelo adequado. Há, basicamente, dois enfoques utilizados nesta análise:

- no primeiro, a análise é feita no domínio do tempo, isto é, estamos interessados na magnitude dos eventos que ocorrem em determinado instante de tempo e na relação entre as observações em diferentes instantes de tempo. A ferramenta utilizada geralmente é a função de autocorrelação (e certas variações da mesma);

- no segundo, a análise é feita no domínio da frequência, isto é, o interesse está na frequência com que certos eventos ocorrem em determinado período de tempo (movimentos cíclicos, por exemplo). A ferramenta usada geralmente é o espectro (que é uma Transformada de Fourier da função de autocorrelação).

Pode-se dizer que a análise no domínio da frequência é conveniente na procura de características determinísticas da série (periodicidades, por exemplo) enquanto que o domínio do tempo é conveniente na procura de características não determinísticas da série (a modelagem de séries temporais que geralmente é efetuada no domínio do tempo, visa encontrar modelos prováveis para aquela série), embora seja também utilizável para processos determinísticos.

Para se efetuar a análise de séries temporais, devemos também levar em conta o espaçamento de suas observações no tempo. As séries em outras áreas geralmente possuem um espaçamento regular no tempo (Por exemplo: produção diária de um produto em determinada indústria). Como é mais comum a existência de séries igualmente espaçadas, a estatística clássica tem elaborado todo um conjunto de teorias para análise e modelagem deste tipo de série. Só para citarmos algumas, temos no domínio do tempo, os seguintes métodos para modelagem:

- média móvel (ou *moving average*- MA)- Tal método consiste na modelagem de uma série temporal através de um processo linear geral, que é um processo que pode ser representado por uma combinação linear ponderada de uma variável aleatória "ruído branco", no presente e no passado:

$$Z_t = a_t + \varphi_1 a_{t-1} + \varphi_2 a_{t-2} + \dots$$

onde Z_t é a série temporal observada; a_t são os termos de "ruído branco", não observados (o "ruído branco" pode ser entendido como um processo aleatório puro cujas variáveis aleatórias têm média nula, variância constante e são não-correlacionadas; e φ_t são coeficientes)

A fórmula dada por Slutsky (1927) e Wold (1938) para a média móvel, segundo a concepção de processo linear geral é:

$$Z_t = a_t - \Theta_1 a_{t-1} - \Theta_2 a_{t-2} - \dots - \Theta_q a_{t-q}$$

- processo auto-regressivo (AR)- É dado pela seguinte fórmula:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t$$

Aquí, em vez de usarmos termos do tipo ruído no presente e no passado, utilizamos termos baseados nas próprias observações Z_t no presente e no passado (combinação linear dos valores mais recentes), mais um termo ruído a_t que incorpora algo novo na série no tempo t , que não foi explicado pelos valores passados (a_t é independente de Z_{t-1}, Z_{t-2}, \dots).

No domínio da frequência geralmente encaramos a série temporal como o resultado de uma função que pode ser expressa por uma soma de senos e cossenos (através da Série de Fourier). A análise da série temporal (neste caso, a verificação de frequências presentes na série) pode ser desenvolvida pela relação entre a estimação de parâmetros da Série de Fourier que representa a série temporal (amplitudes, por exemplo) e a função de autocorrelação. Os parâmetros de Série de Fourier estão intimamente ligados à uma quantidade denominada "Potência Média" que assume valores altos para frequências que estejam presentes na série temporal, isto é, fazemos um gráfico Potência Média vs. frequência (gráfico esse também chamado de espectro) e valores máximos (picos) no gráfico para certas frequências indicam que aquelas frequências estão presentes nos dados da série temporal. Sabe-se que os parâmetros da Série de Fourier (e, portanto, a Potência Média) são Transformadas de Fourier da função de autocorrelação, daí o uso da Transformada de Fourier na análise espectral (domínio da frequência). Uma de suas aplicações é a rotina *Fast Fourier Transform* que,

comparada às outras rotinas de Transformada de Fourier, necessita de menos tempo de computação.

Como vimos, existem muitos procedimentos para análise e modelagem de séries temporais para dados igualmente espaçados. Só que novamente em Astronomia surge um obstáculo: as séries temporais astronômicas geralmente não possuem espaçamento regular de suas observações no tempo, isto é, elas são desigualmente espaçadas. Isso pode ser facilmente entendido: para se observar por exemplo, uma estrela variável, o astrônomo terá que lidar com os seguintes fatores: disponibilidade de tempo do telescópio a ser utilizado; disponibilidade de boas condições meteorológicas; a duração da noite; possíveis problemas na instrumentação assim como o próprio cansaço físico com o passar da noite. Com todos estes fatores, torna-se claro e bem óbvio que dificilmente conseguiremos uma série temporal astronômica igualmente espaçada. Além disso, a estatística clássica não possui uma formulação teórica elaborada para a análise de séries desigualmente espaçadas. É importante, portanto, o desenvolvimento de técnicas e métodos de análise de séries temporais com espaçamento irregular no tempo para que se possa entender e prever o comportamento do sistema estudado. Uma das formas de se estudar este comportamento é verificando se a série contém periodicidades ou não. Este trabalho, portanto, vai se propor a estudar métodos de análise de séries temporais desigualmente espaçadas existentes na literatura, no que se refere a determinação de períodos.

2- BUSCA DE PERIODICIDADE- DESCRIÇÃO DE MÉTODOS DE ESTUDOS DE SÉRIES TEMPORAIS PARA DADOS DESIGUALMENTE ESPAÇADOS.

Como foi dito acima, a determinação de periodicidades em séries temporais astronômicas é uma forma de se prever o comportamento do sistema estudado. Nesta seção, citaremos inicialmente alguns exemplos de sistemas variáveis em Astronomia que contenham periodicidades e faremos um estudo geral dos procedimentos utilizados nos métodos de determinação de períodos. A seguir, estudaremos superficialmente algumas técnicas de determinação de períodos e, finalmente, complicações que podem surgir na aplicação destes métodos devido às características da série.

2.1- Sistemas Variáveis Com Periodicidade:

Em Astronomia, encontramos vários exemplos de sistemas variáveis com periodicidades. Apenas para citarmos alguns exemplos, podemos mencionar as estrelas variáveis intrínsecas, cuja magnitude varia periodicamente com o tempo (em alguns casos devido à pulsação radial da estrela, isto é, contração e dilatação de seu raio); as estrelas binárias eclipsantes, onde temos um par binário e uma estrela brilha menos que a outra. Este sistema possui uma rotação (tal como a rotação de um halteres onde os "pesos" seriam as estrelas) e quando a estrela menos brilhante estiver à frente da mais brilhante na nossa linha de visada, haverá uma queda na luminosidade do sistema, sendo que este fenômeno aparecerá na série temporal da binária eclipsante. Se calcularmos o período desta variação de luminosidade causada pelo "eclipse" de uma estrela em outra, já teremos um parâmetro para estudar o comportamento mecânico (a rotação) do sistema; a rotação de asteróides, onde as variações

periódicas em magnitude dão o período de rotação deste; variáveis cataclísmicas com discos de acreção, onde temos um sistema binário formado por uma anã branca e uma estrela da sequência principal, que perde matéria para a anã branca na forma de um disco de plasma que se forma em torno da anã branca. Partes deste disco são mais quentes e luminosas e podemos determinar o período de rotação do sistema pela variação de luminosidade do disco; o próprio ciclo de atividade solar também é um exemplo de sistema periódico, onde o número de manchas solares, que indicam a atividade solar, varia periodicamente no tempo. Pode-se ver que a Astronomia possui realmente vários tipos de sistemas variáveis periódicos, e seu estudo e análise são bem importantes.

A discussão superficial dos tipos de métodos de determinação de períodos foi tratada com muita clareza por Fullerton (1987) e, devido à sua boa qualidade, nos basearemos nela na discussão que se segue.

Para que possamos estudar sistemas variáveis com periodicidade, devemos primeiro verificar que características possuem suas séries temporais. Das explicações vistas na introdução, já sabemos que uma série temporal astronômica constitui-se de um conjunto de observações X_i de um fenômeno em instantes de tempo t_i , onde i varia de 1 a N . O espaçamento em t_i é geralmente irregular e existe dependência entre as observações vizinhas. Cada observação X_i pode ser encarada em aplicações astronômicas frequentemente como um valor diferencial de magnitude de um objeto num filtro particular ou a velocidade radial de um objeto. Cada observação é contaminada por alguma incerteza aleatória (ruído) devida a erros de medição, flutuações atmosféricas, erros de instrumento, etc. Esta variável aleatória geralmente é assumida como tendo uma Distribuição Gaussiana caracterizada por uma média

nula e um desvio padrão σ . Este tipo de variável aleatória é denominada ruído branco e seus pontos vizinhos não são dependentes, ao contrário dos pontos da série que contém a informação. Aliás, o ruído branco é tido como um processo aleatório puro que não contém informação.

Dadas essas considerações preliminares, comecemos a analisar a busca de periodicidades em sistemas variáveis em linhas gerais. Existem dois objetivos distintos associados com a análise de periodicidades: o primeiro objetivo trata da detecção de períodos, isto é, dada uma série temporal, analisamos se ela possui alguma periodicidade e se esta é estatisticamente significativa. Feito isso, partimos para o segundo objetivo que é a estimação de propriedades do sinal (determinação de períodos, amplitudes), isto é, verificamos quais são os parâmetros associados com o sinal periódico. Estes dois objetivos nem sempre são bem separados pelos pesquisadores e nem sempre é necessário detectar estatisticamente a característica periódica do sinal. Por exemplo, se a série temporal é de uma Cefeida, não precisamos detectar a periodicidade (já sabemos que as Cefeidas são estrelas variáveis com uma periodicidade bem definida na curva de luz).

A metodologia para a determinação de períodos em séries temporais geralmente é da seguinte forma: os dados da série devem ser manipulados e testados por uma função que seja "sensível" a períodos. Criaremos um conjunto de períodos (ou frequências) tentativa e aplicaremos seus valores na função "sensível" a períodos, obtendo vários valores para esta. A função irá atingir um valor extremo (ou um valor máximo ou um valor mínimo, dependendo da natureza do algoritmo) quando o período tentativa for aquele que estiver o mais próximo possível do período da série temporal. Esta inspeção pode ser feita via a geração de um

gráfico com a função sensível a períodos no eixo das ordenadas e o intervalo de períodos tentativa no eixo das abscissas. Tal gráfico é chamado "periodograma", onde o mínimo (ou o máximo) da curva deste gráfico apontará o período tido como o da série temporal (só para recordar, na introdução vimos o conceito de sistema dinâmico que era constituído por uma série de entrada $X(t)$, uma função de transferência $v(t)$ e uma série de saída $Y(t)$. No nosso caso, a série de entrada é a série que queremos determinar o período, a função de transferência é a função sensível a períodos e a série de saída são os pontos do periodograma). Uma vez obtido o periodograma, uma análise posterior é requerida para determinar se os extremos da função são estatisticamente significantes (testes em estatística clássica ou tratamento de *alias* são utilizados geralmente para isso). Convém ressaltar que todos os resultados de métodos de determinação de períodos estão sujeitos a uma variedade de dificuldades originárias da imprecisão inerente às observações, à característica da amostra dos dados e do ruído gerado pelo próprio algoritmo. As propriedades estatísticas dos métodos oferecem a única proteção contra a detecção de períodos falsos.

As características do intervalo de períodos tentativa usado para a resolução da função sensível a períodos depende do propósito da análise. É bem importante que os períodos tentativa do intervalo sejam independentes uns dos outros para que se assegure uma melhor validade do teste estatístico que vai avaliar a significância dos extremos. O intervalo de períodos tentativa deve ser criado de seguinte forma: uma estimativa inicial do período, feita geralmente por inspeção visual da série deve ser realizada. Constrói-se, então, o intervalo de períodos tentativa com o período estimado no centro do intervalo. Este intervalo também deve obedecer o seguinte caso limite: o período mais curto (e frequência mais alta, portanto) é

determinado por intervalos de medidas sucessivas. No caso de dados igualmente espaçados, 3 pontos são suficientes para determinar uma curva senoidal e o período mínimo detectado será $2\Delta t$. A frequência correspondente será conhecida como frequência Nyquist ($\nu_{Ny} = (2\Delta t)^{-1}$).

Contudo, para dados desigualmente espaçados, a frequência Nyquist não pode ser bem determinada (os Δt 's são diferentes), mas podemos definir uma pseudo-frequência Nyquist fazendo a média dos intervalos $\Delta t = t_i - t_{i-1}$ dos dados desigualmente espaçados que é:

$$\nu_{Ny} = (2 \langle \Delta t \rangle)^{-1}.$$

Outro caso limite ao qual o intervalo de períodos tentativa deve obedecer é o mais longo período (frequência mais curta), que é T , isto é, o tempo total que a série temporal possui. A frequência mínima será dada por $\nu_{min} = (T)^{-1}$.

Os casos limites mencionados acima devem ser utilizados para garantir uma resolução mínima do periodograma (o extremo da função sensível a períodos pode não estar neste intervalo, por isso os casos limites geram somente um intervalo mínimo para a análise).

2.2- Métodos para determinação de períodos: um resumo.

Os métodos para determinação de períodos em séries temporais desigualmente espaçadas que vão ser brevemente discutidos nesta seção são aplicáveis somente em séries temporais de curto período (no caso, da ordem de fração do dia) e são divididos em três classes: métodos *string-length*, *phase-binning* e técnicas de Transformada de Fourier (e similares, tais como ajuste por mínimos quadrados). Vamos a elas.

a) Métodos *String-Length* (ou método do tamanho de linha)- A idéia principal deste tipo de método é minimizar o comprimento dos segmentos de reta (*string*) que unem observações

adjacentes no espaço de fase. Fazendo isso, "suavizaremos" ao máximo a curva. Para um período tentativa p as fases ϕ dos dados de entrada são calculadas pela seguinte fórmula:

$$\phi = \text{frac} [(t_i - t_0)/p]$$

onde t_0 é a época adotada e frac indica que somente a parte fracionária da fase será utilizada.

A seguir, as observações são postas em ordem crescente de fase dentro do espaço de fase associado com o período tentativa e o *string-length* é calculado de acordo com a fórmula dada por Dworetzky (1983):

$$l(p) = \sum_n [(f_j - f_{j-1})^2 + (\phi_j - \phi_{j-1})^2]^{1/2} + [(f_1 - f_n)^2 + (\phi_1 - \phi_n + 1)^2]^{1/2} \quad \text{onde } j= 1, 2, \dots, n$$

Obviamente, o *string-length* é calculado por aplicação repetida do Teorema de Pitágoras no espaço de fase. A função $l(p)$ é calculada para um intervalo completo de períodos tentativa e plotada contra p . O menor valor de $l(p)$ ocorre para o período verdadeiro, quando todas as observações caem ao longo de uma curva suave.

Temos duas variações do método dado por Dworetzky (1983). O método Lafler e Kinman (1964) foi o primeiro a utilizar computadores. Entretanto, ele somente calcula a componente vertical do *string-length*, isto é, a distância vertical entre os pontos adjacentes. Desprezar a componente horizontal previne uma certa tendência em regiões povoadas da amostra mas aumenta o ruído no periodograma. A técnica de Renson (1978) inclui uma função peso que permite a análise para uma distribuição desigual de observações no espaço de fase.

A vantagem básica deste método é a sua simplicidade. A desvantagem que o método sofre é o fato de que para cada período tentativa as observações, com suas respectivas fases, devem ser dispostas em ordem crescente de fase. Isso aumenta o tempo de computação.

b) Métodos *phase-binning* (ou método de armazenamento de fases)- Técnicas que separam os dados entrada utilizando uma frequência ou período tentativa e examinam os conteúdos de "caixas" em suas fases são frequentemente utilizados em estudos astronômicos. Métodos desse tipo são os descritos por Jurkevich (1971) e Stellingwerf (1978).

O método descrito por Stellingwerf é, em linhas gerais, da seguinte forma: dada uma série temporal, a variância da população (isto é, de todos os pontos da série temporal é dada pela seguinte fórmula:

$$\sigma^2 = \sum_{i=1}^N (x_i - \bar{x})^2 / (N - 1)$$

onde x_i é a observação no instante t_i , e \bar{x} é a média das observações. Vamos escolher m amostras desta população de N observações. Cada amostra terá uma variância S_j^2 , $j=1, 2, \dots, m$ e conterá n_j pontos. A variância total (unindo-se a variância de todas as amostras) é dada pela fórmula:

$$S^2 = \sum_{j=1}^m (n_j - 1) S_j^2 / \sum_{j=1}^m (n_j - m)$$

Seja, agora, Π o período tentativa e ϕ a fase. Temos que ϕ é:

$$\phi_i = \text{FRAC}(t_i / \Pi)$$

que é a função $t \text{ MOD } \Pi$, isto é, ϕ só assume valores entre zero e um. Vamos usar ϕ como o critério de escolha das m amostras. Cada x_i terá o seu ϕ_i e cada amostra conterá pontos x_i que

tenham ϕ_i similares. A variância dessas amostras dará uma medida do espalhamento em torno da curva de luz média definida por meio de $x_i(\phi)$ em cada amostra.

Definiremos, então, a estatística (ou função sensível a períodos):

$$\Theta = S^2 / \sigma^2$$

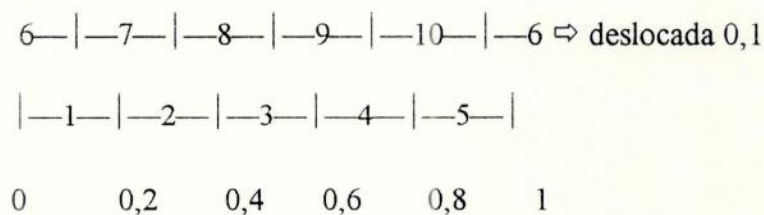
onde S^2 é a soma das m variâncias e σ^2 é a variância da população.

- Se o período tentativa se aproxima daquele da série temporal, $S^2 \rightarrow 0$ (dispersão mínima entre os pontos) e Θ alcançará um valor mínimo.
- Se o período tentativa diferir muito daquele da série temporal, $S^2 \rightarrow \sigma^2$ (dispersão alta).

Plotando um gráfico Θ vs. período tentativa, os mínimos da curva apontarão os períodos tomados como os da série.

Testes estatísticos utilizados para se verificar a significância dos mínimos são o teste F e o χ^2 (para o caso de $N > 100$).

Foi dito que, no critério de escolha das m amostras, estas conteriam pontos x_i com ϕ_i similares. O critério de similaridade será explicado pela estrutura de caixa utilizada neste método. Esta pode ser entendida da seguinte forma: dado o espaço de fase (intervalo entre 0 e 1), dividimos este em, por exemplo, 5 partes e chamaremos estas partes de grupos ou "caixas". Cada caixa terá tamanho 0,2. Haverá outra estrutura de caixa acima desta (do espaço de fase) deslocada de 0,1, como mostra a figura abaixo:



Temos, portanto, 10 caixas e cada observação vai cair em duas caixas de acordo com sua fase. Por exemplo: se a observação x_i possuir fase 0,08 (calculada por $t \text{ MOD } \Pi$), x_i vai cair na caixa 1 e na caixa 6. Se x_i possuir fase 0,12, este vai cair nas caixas 1 e 7. Assim, são geradas as $m = 10$ amostras e aplica-se o método. Esta estrutura de caixa pode ser denotada por (N_b, N_c) onde N_b é o número de caixas no espaço de fase e N_c é o número de caixas onde cai a observação. $1/(N_b N_c)$ nos dá o deslocamento da estrutura de caixa superior e $N_b N_c$ nos dá o número total de caixas.

Ainda nos métodos *phase-binning*, analisaremos agora o método descrito por Jurkevich (1971). Dada uma série temporal, faremos estimativas para o seu período, quer dizer, imaginaremos séries "hipotéticas" cujos períodos serão os períodos tentativa. Compararemos cada série com a série original e verificaremos o quanto a série do período tentativa irá "dispersar" (não "encaixar") em relação à série original. A que melhor se "encaixar" (tiver a menor dispersão), terá o seu período tentativa apontado como o verdadeiro. Para isso, novamente utilizaremos a fase ϕ como:

$$\phi_i = \text{FRAC}(t_i / \Pi) = t \text{ MOD } \Pi \quad 0 < \phi_i < 1$$

e subdividiremos este intervalo em vários grupos ou "caixas". Cada fase estará associada a uma observação. Dividiremos a população em m amostras, onde m é o número de grupos de acordo com o seguinte critério: cada amostra conterá as observações X_i que possuam fases que estejam no intervalo do grupo (Por exemplo: dividimos o intervalo entre zero e um das fases em cinco grupos: o primeiro grupo vai de 0 a 0,19; todas as observações que tenham fases entre 0 e 0,19 serão postas naquela primeira amostra. O segundo grupo vai de 0,2 a 0,39; todas as observações que tenham fases entre 0,2 e 0,39 serão postas naquela segunda

amostra; e assim sucessivamente). Calculamos o desvio médio quadrático para cada amostra e o somatório dos desvios médios quadráticos será a estatística V_m^2 (ou função sensível ao período). Plotamos V_m^2 vs. período tentativa e o mínimo desta curva estará apontando o período tido como o verdadeiro da série original. Testes estatísticos utilizáveis são o teste F (embora com muitas restrições) e a estatística de Bartlett.

Assim como os métodos *string-length*, os métodos *phase-binning* não fazem afirmações sobre a forma da variação estudada. Contudo, métodos *phase-binning* não são tão computacionalmente intensivos quanto os *string-length*, principalmente devido ao fato de que os dados não precisam ser ordenados dentro dos grupos ou "caixas".

c) Transformada de Fourier, técnicas de periodograma e mínimos quadrados- A Transformada de Fourier e técnicas relacionadas são todas baseadas na *Discrete Fourier Transform* e empregam conceitos elegantes incorporados em análise harmônica. Embora aparentemente distintos, todos estes métodos de detecção de períodos são conceitualmente similares. A Transformada de Fourier e técnicas de periodograma estão intimamente relacionadas sendo a última essencialmente o módulo quadrado da primeira adequadamente normalizada (Deeming, 1975; Scargle, 1982). Técnicas de ajuste de mínimos quadrados são formalmente equivalentes a análise de periodogramas sob certas circunstâncias.

A popularidade destas técnicas é devida a eficiência computacional da rotina da *Fast Fourier Transform*. Infelizmente, este algoritmo requer dados de entrada que estejam igualmente espaçados no tempo, o que é praticamente impossível em Astronomia.

Uma formulação extremamente usual da função periodograma foi desenvolvida por Scargle (1982):

$$P_X(\omega) = \frac{1}{2}[C^2(\omega) + S^2(\omega)]$$

onde:

$$C(\omega) = A(\omega) \sum X(t_j) \cos(\omega|t_j - \tau|)$$

$$S(\omega) = B(\omega) \sum X(t_j) \text{sen}(\omega|t_j - \tau|)$$

$$A(\omega) = (\sum \cos^2(\omega|t_j - \tau|))^{1/2}$$

$$B(\omega) = (\sum \text{sen}^2(\omega|t_j - \tau|))^{1/2}$$

e onde $\tau(\omega)$ é definido por:

$$\text{tg}(2\omega\tau) = (\sum \text{sen}(2\omega t_j)) / (\sum \cos(2\omega t_j))$$

Nestas expressões, ω é a frequência angular, $\omega = 2\pi\nu = 2\pi(P)^{-1}$, X_j são os dados entrada obtidos em tempos espaçados arbitrariamente t_j , $j=1$ a N e $\tau(\omega)$ é uma função fase designada para fazer este periodograma formalmente equivalente às técnicas de ajuste de mínimos quadrados. Para entender como esta função é sensível à frequências (ou períodos) presentes nos dados, suponhamos que o conjunto de dados X_j contenha um sinal senoidal de frequência angular ω_0 . Quando a frequência tentativa não é próxima de ω_0 , os dados entrada X_j e os termos $\cos(\omega|t_j - \tau|)$ e $\text{sen}(\omega|t_j - \tau|)$ estão fora de fase. Os termos nos somatórios possuem sinais distribuídos aleatoriamente com respeito a cada um deles e, portanto, se cancelam. Contudo, quando $\omega \cong \omega_0$, os dados e os fatores senoidais estão aproximadamente em fase e os termos nos somatórios adicionam-se coerentemente tal que $P_X(\omega)$ atinge um valor relativamente grande (máximo). Assim, o periodograma revela periodicidades escondidas por "ressonância".

O cálculo do periodograma de Scargle é equivalente a um ajuste de senoidais aos dados pelo método de mínimos quadrados linear. Exemplos deste tipo de método são dados por Vanicek (1969) e Ponman (1981). Para cada uma das frequências tentativa, uma combinação linear de senos e cossenos com argumento $(\omega t - j)$ são ajustados aos dados por mínimos quadrados. O somatório das amplitudes quadradas de ambos estes termos (seno e cosseno) é plotada contra a frequência para formar um gráfico que se assemelhe a um periodograma. Picos no gráfico indicam períodos; as alturas dos picos indicam significâncias através de sua relação com a estatística F.

As técnicas de Fourier e suas similares diferem das técnicas *phase-binning* e *string-length* pelo fato de que o sinal é assumido como tendo uma forma funcional definida (senoidal). Isto parece que vai limitar a utilidade destas técnicas para a análise de variações não-senoidais. Em princípio, os métodos de Fourier são aplicáveis à ondas de forma arbitrária pelo Teorema de Decomposição de Fourier, que estabelece que qualquer sinal periódico pode ser decomposto em um somatório de senoidais (possivelmente infinito). A frequência de cada uma destas senoidais será um múltiplo inteiro da frequência fundamental ν_0 , por exemplo, $2\nu_0, 3\nu_0, 4\nu_0$, etc. Isso ocorre para curvas de luz de Cefeidas e Binárias Eclipsantes.

2.3- Complicações:

Infelizmente, todos os procedimentos de análise de períodos sofrem várias complicações. Elas podem vir sobre várias circunstâncias.

a) *Aliasing*.

Este fenômeno inibe e às vezes evita a identificação da frequência correta. Ele se origina quando os dados entrada são distribuídos em intervalos igualmente espaçados. Esta regularidade permite que sinais de frequências $\nu_0 \pm n(\Delta t)^{-1}$ se ajustem aos dados igualmente bem, causando grande confusão. Como as séries temporais astronômicas são geralmente desigualmente espaçadas, o *aliasing* não se torna tão preocupante.

Contudo, um fenômeno similar denominado *pseudo-aliasing* está quase sempre presente em observações astronômicas. Ele causa a mesma confusão que o *aliasing* quando os dados possuem espaços em branco (lacunas) de forma regular, quer dizer, existe uma regularidade no padrão das observações. Por exemplo, periodogramas de observações em noites consecutivas mostrarão fortes *pseudo-aliases* de $\pm dia^{-1}$. Formas de se evitá-los seriam: fazer observações de fenômenos com variação rápida (mensurável em uma noite); conseguir tempo num telescópio que estivesse nos pólos em seu inverno; uma rede de vários telescópios ligados via satélite a diferentes longitudes.

Existem, entretanto, alguns procedimentos para se tratar *pseudo-aliases*. Em técnicas de Fourier, podemos calcular uma "função janela" (isto é, um periodograma de variável independente). Uma comparação entre os picos da função janela com o padrão de *pseudo-aliases* no periodograma pode ajudar na identificação de frequência verdadeira que é geralmente, embora nem sempre, o maior pico do periodograma.

b) Estimativa de ruído.

A única proteção contra detecção de períodos falsos cai numa rigorosa análise estatística. Uma estimativa acurada do nível de ruído presente nos dados é de considerável

importância uma vez que a relação sinal-ruído é a quantidade pela qual a significância do pico do periodograma é acessada.

Existe uma variedade de enfoques para estimativa de ruído num conjunto de dados. A precisão externa das observações (Por exemplo: Desvio-Padrão de observações em fotometria ou o erro médio de uma medida de velocidade radial) providencia uma estimativa direta. Em princípio, o ruído pode ser determinado diretamente de um periodograma tipo Fourier sob a afirmação de que o ruído é branco localmente. Esta forte afirmação (ruído branco localmente) é necessária porque observações fotométricas sofrem flutuações em baixas frequências devido à mudanças horárias e noite a noite nas condições atmosféricas.

c) *Prewhitening* e a detecção de múltiplos períodos.

Prewhitening é um termo usado para indicar a remoção de um sinal para uma frequência particular do conjunto de dados. Usualmente este é completado por ajuste de senoidais de frequência apropriada aos dados por mínimos quadrados. Os resíduos do ajuste formam os dados *prewhitened*.

Prewhitening tem dois usos em análises de séries temporais. O primeiro é ajudar a identificação de uma (simples) periodicidade "verdadeira" sob condições desfavoráveis de *pseudo-aliasing*. Para se decidir quais dos picos estatisticamente significantes correspondem à frequência real, os dados são "*prewhitened*" para cada uma das frequências de *alias* e periodogramas dos resíduos são examinados. Quando uma frequência de *pseudo-alias* tiver sido removida, os outros picos vão permanecer desde que um sinal "físico" ainda esteja presente. Frequentemente, o periodograma dos resíduos contém novos picos e o nível de ruído aumenta. Em contraste, fazer o *prewhitening* com a frequência de sinal verdadeiro causa a

remoção de todos os picos de *pseudo-alias* do periodograma dos resíduos; o ruído total usualmente decresce. Assim, a frequência verdadeira pode ser distinguida do *pseudo-alias* em muitos casos.

Muito frequentemente, o *prewhitening* é usado para pesquisar conjuntos de dados com mais de uma periodicidade. Este procedimento envolve a detecção de um período através da análise de periodograma, sua subsequente remoção através do *prewhitening* e uma reanálise dos dados residuais para mais períodos. Em princípio, este enfoque passo a passo é válido para uma análise de uma combinação linear de sinais senoidais multiperíódicos. Na prática, a detecção de mais de um período é muito difícil, primariamente porque o *prewhitening* destrói a independência dos dados entrada. Se isso ocorre, os ajustes de *prewhitening* possuirão grandes incertezas. Como resultado, as significâncias estatísticas dos picos do periodograma de resíduos que sofrem *prewhitening* são impossíveis de serem avaliadas estatisticamente. Frequentemente, não mais do que uma periodicidade pode ser extraída com segurança do conjunto de dados.

3- ALGORITMOS PARA MÉTODOS SELECIONADOS:

Após uma breve apresentação das técnicas para determinação de periodicidades em sistemas variáveis de curto período, iremos agora estudar de forma mais aprofundada duas destas técnicas. Veremos o método tipo *phase-binning* descrito por Jurkevich (1971) e um método de ajuste por mínimos quadrados (Ponman-1981). Será feita uma descrição dos

métodos e dos programas elaborados para os mesmos. Começemos, então com o método descrito por Jurkevich (1971).

3.1- Método Jurkevich.

a) Descrição do método.

Para um melhor entendimento do método, façamos a sua descrição num contexto inicialmente qualitativo: dada a série temporal, faremos "estimativas" para o seu período (intervalo de períodos tentativa). Para isso criaremos "séries temporais hipotéticas" cujos períodos serão as estimativas para a série temporal que analisamos e para a qual queremos determinar o período (a chamaremos daqui por diante de série original). Vamos comparar cada série hipotética com a original e verificamos o quanto a série do período tentativa irá "não encaixar" (isto é, dispersar) da série original. A que melhor se "encaixar" na série original terá o seu período tentativa apontado como o verdadeiro da série original. As dispersões entre os pontos da série hipotética e a original serão medidos com o auxílio da diferença de fase entre eles.

Desenvolvamos estas idéias agora quantitativamente: dados os pontos da série original (X_i) com seus respectivos t_i , as fases ponto a ponto entre a série original e a série do período tentativa são calculadas pela seguinte fórmula:

$$\phi_i = \text{FRAC}(t_i / \text{período tentativa}) = t \text{ MOD período tentativa} \quad (i=1, 2, \dots, N)$$

N)

Observação: a fase só poderá estar entre 0 e 1 (caso seja maior que 1, pegamos somente sua parte fracionária). Teremos, então, N fases de valores entre 0 e 1, onde o valor 1 é arbitrado como 0 (função MOD). O intervalo entre 0 e 1 será subdividido (de forma fictícia, abstrata) em vários grupos ou "caixas". Por exemplo, se dividirmos este intervalo em 5 caixas, a primeira caixa teria o intervalo [0 ; 0,19]; a segunda caixa, [0,2 ; 0,39]; a terceira caixa, [0,4 ; 0,59]; a quarta caixa, [0,6 ; 0,79]; e a quinta caixa, [0,8 ; 0,99]. Vamos ao conjunto de N fases e as separamos, "jogando" em cada caixa as fases com valores naquele intervalo (se a fase tem valor 0,13, cai na caixa 1; se tem valor 0,36, cai na caixa 2; se tem valor 0,87, cai na caixa 5 e repete-se todo o procedimento para todas as outras fases). Teremos, conseqüentemente, 5 conjuntos de fases. Cada conjunto destes terá m_l fases onde $l=1, 2, \dots, 5$ (neste caso) e cada fase estará associada a um ponto da série original X_i (isto é, a uma magnitude). Portanto, na verdade, teremos cinco conjuntos de magnitudes X_i que foram organizados de acordo com o critério imposto acima pelas fases. Calcularemos o \overline{X}_l que é a magnitude média para cada caixa:

$$\sum_{i=1}^{m_l} X_i / m_l = \overline{X}_l$$

e o desvio médio quadrático para cada caixa:

$$V_l^2 = \left[\sum_{i=1}^{m_l} X_i^2 \right] - m_l [\overline{X}_l]^2$$

Além disso, calcularemos V_m^2 (soma dos desvios médios quadráticos de todos os grupos ou caixas):

$$V_m^2 = \sum_{l=1}^m V_l^2 \quad (m=5, \text{ no nosso exemplo})$$

Pelo teorema de adição de variâncias, temos geralmente $V_m^2 < V^2$ (onde V^2 é a variância da amostra de X_i sem subdividi-la em caixas, isto é, a variância da população); como consequência deste teorema, se os V_i^2 forem diferentes, temos $V_m^2 < V^2$ e se os V_i^2 forem iguais, $V_m^2 \cong V^2$. A diferença $V_{BG}^2 \cong V^2 - V_m^2$ mede então se a dispersão é alta ou não. Se a curva da série hipotética com o período tentativa não se adequar bem à da série original, teremos alta dispersão para as fases das caixas e suas variâncias serão também aproximadamente as mesmas (pequeno V_{BG}^2 , grande V_m^2); se a série do período tentativa se encaixar bem à da série original, teremos V_i^2 diferentes (grande V_{BG}^2 , pequeno V_m^2). O período tentativa que tiver o menor V_m^2 será o período verdadeiro (isto é plausível, pois a menor dispersão total entre a série observada e a do período tentativa mostra que elas têm aproximadamente o mesmo período). Como temos vários períodos tentativa, repetimos todo o procedimento descrito acima e faremos um conjunto de V_m^2 , sendo que cada V_m^2 será atribuído para um período tentativa. Se plotarmos V_m^2 vs. período tentativa, o período real será estimado a partir da análise do mínimo da curva (cabe observar que o período tentativa estimado estaria no meio do intervalo de períodos tentativa utilizado).

Um teste estatístico utilizado para se verificar a significância dos resultados do método é o teste F, embora com algumas restrições:

- os erros observacionais devem ser assumidos como normalmente distribuídos;
- todos os grupos necessitam conter observações;
- Jurkevich verificou a condição acima satisfeita para $N=200$ e $m=10$ (embora isto não necessariamente ocorra somente para este caso).

O teste estatístico utilizado neste trabalho será o teste Bartlett onde testamos a hipótese de que as variâncias das m caixas são iguais. Se isso ocorrer, não haverá significância para aquele V_m^2 (lembrar que se os V_i^2 forem iguais, $V_m^2 \cong V^2$, implicando num grande V_m^2), isto é, ele não será estatisticamente significante. Assim, num periodograma de uma série monoperiódica com mais de um mínimo, teremos condições de avaliar qual dos mínimos é o significante. O teste é dado pela seguinte fórmula:

$$M = (N - K) \ln \left[\frac{\sum_{i=1}^m (N_i - 1) S_i^2}{(N - K)} \right] - \sum_{i=1}^m [(N_i - 1) \ln S_i^2]$$

onde N: número de dados da população

K: número de caixas (ou o m)

N_i : número de dados nas caixas (ou os m_i)

S_i^2 : variância em cada caixa (ou os V_i^2)

Calcula-se M para cada V_m^2 e com o auxílio da Tabela χ^2 , para um nível de significância de 5%, verificamos a condição para (K-1) graus de liberdade:

$$M < \chi^2$$

(Hamilton, 1964). Se a hipótese de que as variâncias das caixas sejam iguais for rejeitada, o respectivo V_m^2 terá embasamento estatístico para se tornar confiável.

b) O programa.

Foi elaborado um programa (em PASCAL) para aplicar o método. Ele contém primeiro um bloco que lê um arquivo em ASCII, arquivo esse que contém os dados da série, além de possuir um contador para os mesmos. A segunda parte do programa faz uma leitura do número de caixas desejado (no intervalo de 0 a 1 das fases) e do intervalo de períodos

tentativa (com escolha de precisão). A seguir, temos o bloco principal com todas as subrotinas necessárias para cálculos de período tentativa, fases, agrupamento de fases em caixas, cálculos de $m_l, \overline{X}_l, V_l^2$ para as caixas e, finalmente, de V_m^2 para cada período tentativa, além do cálculo da Estatística de Bartlett (as operações são repetidas recursivamente para cada período tentativa). Continuando, temos um bloco que mostra na tela os valores de V_m^2 , de seus respectivos períodos tentativa e da estatística de Bartlett para cada V_m^2 e outro bloco que mostra, no final, o período verdadeiro (aquele que possui o menor V_m^2 dentre aqueles do conjunto de dados). O último bloco consiste no armazenamento destes dados em arquivo. O programa ainda contém um artifício para retomarmos os cálculos e utilizarmos uma precisão maior (no intervalo de períodos tentativa). A escolha da precisão do intervalo corresponde ao número de casas decimais. Por exemplo, se quisermos um intervalo de períodos tentativa de 0,20, 0,21, 0,22, ..., até 0,35, devemos colocar como limites inferior e superior os valores 0,20 e 0,35, respectivamente. A seguir, na ordem de grandeza das subdivisões, o valor 100. Se quisermos agora o intervalo 0,200, 0,201, 0,202, ..., até 0,350, fazemos 0,20 e 0,35 para limites inferior e superior e ordem de grandeza 1000.

3.2- Método Ponman (Mínimos quadrados):

Este método é uma versão, para dados desigualmente espaçados, das técnicas de Fourier (como já foi dito, técnicas de Fourier tipo *Fast Fourier Transform* e outras só são eficientes para dados com espaçamento regular no tempo). Ele consiste no mais simples ajuste por mínimos quadrados. Cada componente seno e cosseno do espectro é ajustado individualmente aos pontos dos dados. Isto simplifica bastante a execução da técnica, embora

um refinamento seja necessário já que o método gera resultados aproximados devidos à estas simplificações.

a) Descrição do método:

Vamos usar, para simplicidade, a seguinte notação:

$$\begin{aligned}
 c(t) &\equiv \cos(2\pi\nu t) & s(t) &\equiv \text{sen}(2\pi\nu t) \\
 c_n &\equiv \cos(2\pi\nu t_n) & s_n &\equiv \text{sen}(2\pi\nu t_n) \\
 c'_n &\equiv \cos(2\pi\nu' t_n) & s'_n &\equiv \text{sen}(2\pi\nu' t_n) \\
 c_n^{(m)} &\equiv \cos(2\pi\nu^{(m)} t_n) & s_n^{(m)} &\equiv \text{sen}(2\pi\nu^{(m)} t_n) \\
 f_n &\equiv f(t_n)
 \end{aligned}$$

Os dados serão sempre indicados pelo índice n que pode tomar valores de 1 a N (onde N é o número total de dados).

Façamos, agora, a descrição do método propriamente dito. Suponhamos que os dados da série temporal possam ser representados por uma função que seja a soma de um termo de ruído aleatório (o qual inclui tanto flutuações aleatórias na saída da fonte quanto flutuações devidas ao processo de recepção) com um sinal periódico de frequência ν' :

$$f(t) = \varepsilon(t) + \sum_{m=1}^{\infty} [\alpha_m c^{(m)}(t) + \beta_m s^{(m)}(t)]$$

onde $c^{(m)}(t) = \cos(2\pi\nu^{(m)}t)$ e $\nu^{(m)} = m\nu'$. Vamos assumir que a média dos dados e qualquer tendência lenta sejam subtraídas tal que $\varepsilon(t)$ é uma flutuação aleatória em torno de zero. Nossos dados consistem, na verdade de valores $f(t_n)$ para $n=1, 2, \dots, N$, onde N é o número total de dados. Destes dados desejamos estimar ν' e os valores dos coeficientes α_m e β_m (isto é, as amplitudes dos termos seno e cosseno).

Vamos considerar a seguinte expressão para um ν qualquer:

$$A(\nu) = \sum_n f_n c_n / \sum_n c_n^2$$

usando a fórmula para $f(t)$, temos:

$$\begin{aligned} A(\nu) &= (1 / \sum_n c_n^2) \sum_n \{ \varepsilon_n + \sum_{m=1}^{\infty} [\alpha_m c^{(m)}(t) + \beta_m s^{(m)}(t)] c_n \} = \\ &= (1 / \sum_n c_n^2) \{ \sum_n \varepsilon_n c_n + \sum_n \alpha_1 c_n' c_n + \sum_n \alpha_2 c_n'' c_n + \dots + \sum_n \beta_1 s_n' c_n + \dots \} \end{aligned}$$

onde $\varepsilon_n \equiv \varepsilon(t_n)$. Assumimos que o número de pontos é aproximadamente grande ($N \geq 20$, Ponman 1981) e $1/\nu \ll T$ para $T = t_n - t_1$. T seria o período de uma onda do tamanho da série. Para $1/\nu \ll T$ ocorrer, precisamos de um grande ν , implicando em pequenos comprimentos de onda. Possuímos, então, muitas ondas dentro desta suposta onda do tamanho da série.

Das propriedades de funções ortogonais temos que:

$$\int_{-T/2}^{T/2} \cos(m\omega_0 t) \cos(n\omega_0 t) dt = 0, \text{ se } m \neq n; T/2, \text{ se } m = n.$$

Substituindo esta propriedade nas funções acima, podemos dizer que o somatório \sum_n vai fazer o papel da integral em T ; N assumiria o papel de T para $m=n$ e teríamos:

$$\sum_n c_n^2 \cong N/2$$

O termo de ruído $\varepsilon(t)$ é pequeno comparado com $N/2$ assim como os termos $\sum \cos(a) \cos(b)$ que vai se anular para $a \neq b$ e $\sum \sin(a) \cos(b)$ que vai se anular para $\forall a$ e b (propriedades das funções ortogonais). Assim, geralmente $A(\nu)$ será pequeno. Se, contudo, ν tomar um valor de uma das frequências presentes nos dados (do tipo $\nu^{(m)}$) então o termo $\sum_n \alpha_m (c_n^{(m)}) (c_n^{(m)})$ em $A(\nu)$ não será pequeno. Os outros termos ($\sum \cos(a) \cos(b)$ para $a \neq b$

e $\sum \text{sen}(a)\cos(b)$ para $\forall a$ e b vão se anular e $A(v^{(m)}) \cong \alpha_m$ devido ao fato de que $A(v^{(m)}) = (1/N/2) \{ \sum_n \alpha_m c_n^{(m)} c_n^{(m)} \} = (2/N)(\alpha_m N/2)$. Isso nos diz que $A(v)$ terá um valor pequeno, a menos que v seja a frequência presente nos dados, e neste caso $A(v)$ toma o valor da amplitude dos cossenos.

Similarmente, introduzimos:

$$B(v) = \sum_n f_n s_n / \sum_n s_n^2$$

que reconheceremos como a amplitude do seno.

Assim, se interpretarmos os pontos da série temporal como resultados de uma função $f(t)$ que consiste na soma de um termo ruidoso com uma combinação linear de senos e cossenos, $A(v)$ e $B(v)$ darão as amplitudes do cosseno e do seno (α_m e β_m , respectivamente) para a frequência verdadeira da série. A função "sensível" à frequências, isto é, a estimativa espectral é dada pela fórmula:

$$P(v) = A^2(v) + B^2(v).$$

Podemos levar em consideração as incertezas diferenciais dos dados em cada ponto usando um procedimento estatístico normal. Se σ_n é o desvio-padrão do n -ésimo ponto (como uma série temporal é constituída em cada ponto de um conjunto possível de valores onde só um ocorre, cada ponto pode possuir um desvio-padrão, chamado aqui de σ_n , onde os dados utilizados para o cálculo de σ_n seriam os valores possíveis e o que realmente foi encontrado), então ponderamos os termos em nossos somatórios com o fator $1/\sigma_n^2$, dando:

$$A(v) = \sum_n f_n (c_n / \sigma_n^2) / \sum_n (c_n^2 / \sigma_n^2) \quad \text{e} \quad B(v) = \sum_n f_n (s_n / \sigma_n^2) / \sum_n (s_n^2 / \sigma_n^2)$$

Como o cálculo de σ_n^2 é tecnicamente impossível, neste trabalho aplicaremos em seu lugar os erros de observação.

b) Aproximações e erros.

Existem três razões pelas quais a estimativa espectral pode desviar do espectro de potências verdadeiro da fonte.

Primeiro, os dados observados contêm flutuações estatísticas e erros de observação que estão inseridos no desvio-padrão σ_n de cada ponto.

Hoyng (1976) derivou uma expressão para o desvio-padrão numa estimativa espectral Fourier assumindo erros iguais para cada ponto. Refazendo essa teoria para um caso genérico (embora usemos erros iguais para cada ponto), temos o resultado:

$$\bar{\varepsilon}(\nu) = 2\sqrt{(A^2(\nu)/C_\nu) + (B^2(\nu)/S_\nu) - (1/2C_\nu^2) - (1/2S_\nu^2)}$$

$$\text{onde } C_\nu \equiv \sum_n (c_n^2 / \sigma_n^2) \text{ e } S_\nu \equiv \sum_n (s_n^2 / \sigma_n^2)$$

Este é o desvio-padrão do ponto $P(\nu)$ da estimativa espectral devida ao ruído. Ele permite-nos julgar a significância do maior pico que aparecerá no periodograma, desde que as duas outras fontes de erro abaixo sejam desprezíveis.

A segunda fonte de erros é o cancelamento imperfeito de termos cruzados. Foi dito que termos cruzados tipo $\cos(2\pi\nu t_n)\cos(2\pi\nu t_n)$ se anulam em comparação com $\cos(2\pi\nu t_n)\cos(2\pi\nu t_n)$ quando somados sobre n . Às vezes, termos cruzados não se cancelam e quando isso ocorre, precisamos introduzir na estimativa espectral refinamentos necessários. A tendência e a variância introduzidas na estimativa espectral devido ao cancelamento imperfeito de termos cruzados são dados neste caso por:

$$b \equiv (4 \sum_n (1/\sigma_n^4) / (\sum_n (1/\sigma_n^2))^2) \sigma_f^2 \quad \text{e} \quad \sigma_p^2 \leq (16 \sum_n (1/\sigma_n^4) / (\sum_n (1/\sigma_n^2))^2) \sigma_f^2 P(\nu)$$

onde $\sigma_f^2 = (1/N) \sum_n f_n^2$ é a variância da população. Estas últimas (b e σ_p^2) são maiores que a incerteza produzida pelo ruído nos dados ($\bar{\varepsilon}(\nu)$). Contudo, se o espaçamento nos dados possui qualquer regularidade, então b e σ_p^2 dão superestimativas grosseiras. Por exemplo,

$[\sum_n \cos(2\pi\nu t_n) \cos(2\pi\nu' t_n)]^2$ toma valores $\cong N/4$ se t_n é aleatoriamente distribuído, mas quando cada t_n é restrito a cair dentro de um intervalo T_b centrado num conjunto de valores regularmente espaçados, então quando temos valores abaixo de $1/2\nu^*$ (onde ν^* é o maior valor entre ν e ν') para T_b , os valores de termos cruzados caem de $N/4$ para 1. Portanto, para uma amostra semi-regular, b e σ_p^2 caem por um fator $N/4$.

Uma fórmula aproximada para $\bar{\varepsilon}_p^2(\nu)$ (a variância devida aos erros dos dados) é

$$\bar{\varepsilon}_p^2 \cong 8P(\nu) / \sum_n (1/\sigma_n^2)$$

tal que o erro introduzido pelo não cancelamento de termos cruzados vai ser desprezível comparado ao erro $\bar{\varepsilon}_p^2(\nu)$ (para o caso de uma amostra semi-regular) provendo-nos a desigualdade:

$$(4/N)(16 \sum_n (1/\sigma_n^4) \sigma_f^2) / (\sum_n (1/\sigma_n^2))^2 \ll 8 / \sum_n 1/\sigma_n^2$$

e, rearrumando os termos:

$$N \gg 8\sigma_f^2 \sum_n (1/\sigma_n^4) / \sum_n (1/\sigma_n^2)$$

Se a condição acima não for satisfeita ou a amostra não for semi-regular, então a variância σ_p^2 devida aos termos cruzados, deve ser adicionada à variância de ruído ($\bar{\varepsilon}^2$) para se obter a variância total da estimativa espectral.

A terceira fonte de incertezas é o já citado fenômeno de *aliasing*. Como foi visto, o *alias* é o fenômeno de detecção de frequências (ou períodos) que não são as verdadeiras devido ao igual espaçamento dos dados onde mais de uma frequência se adequa aos pontos. Existem outras formas pelas quais o *alias* ocorre. Como já foi visto, $A(\nu)$ assumiria um valor significativo para as frequências ν', ν'', \dots etc. do sinal formado pelos pontos de $f(t)$ (onde

$\nu^{(m)} = m \nu'$). Se $A(\nu)$ assumir um valor significativo para um ν que seja diferente de $\nu', \nu'', \dots etc.$, podemos dizer que este é um caso de *alias*. Como este *alias* ocorre?

O termo de ruído deve ser pequeno (assumindo novamente média e tendência lenta subtraídas) mas erros e incertezas remanescentes podem ser significantes em 3 formas:

1) Se ν está próxima a uma das frequências presentes nos dados, então ela pode influenciar no ajuste. Não chega a produzir picos espúrios mas causa um alargamento $\Delta \nu \approx 1/T$ nos picos onde T é o intervalo total de tempo da série. Picos satélites podem aparecer, mas podem ser minimizados, além de possuírem menor amplitude.

2) Se os dados consistem de vários blocos extensamente separados, então a onda que vai se ajustar aos dados só vai precisar se ajustar nestes blocos. Isto leva a formação de picos satélites que se estendem sobre um intervalo $\approx 1/T_{m\acute{a}x}$ (onde $T_{m\acute{a}x}$ é o comprimento do mais longo bloco de dados) em torno de qualquer pico verdadeiro e que podem ser de tamanho comparável ao pico principal (também conhecido como fenômeno de *pseudo-alias*).

3) Um fenômeno similar à reflexão em torno da frequência Nyquist, que encontramos num espectro Fourier convencional, ocorre se os dados são igualmente espaçados. Suponha que as medidas sejam tomadas semi-regularmente com uma separação típica Δt . Então, para que o "casamento" entre os termos $\cos(2\pi\nu t)$ e $\cos(2\pi\nu' t)$ gere valores grandes para

$\sum_n \alpha_n c_n' c_n / \sigma_n^2$, devemos ter que $\nu = (q / \Delta t) \pm \nu'$, onde q é um inteiro. Então um pico no periodograma para ν' dá origem a picos de *alias* para valores de ν que satisfaçam a equação $\nu = (q / \Delta t) \pm \nu'$. Este é justamente o mesmo resultado para a análise Fourier padrão onde a frequência Nyquist é $\nu_{Ny} = (2\Delta t)^{-1}$ (lembrar que a frequência Nyquist é a frequência ótima de resolução, isto é, a informação que pode ser obtida de uma amostra com o menor número de dados possíveis; geralmente, para dados igualmente espaçados, são 3 pontos; daí o $(2\Delta t)^{-1}$).

Aqui, contudo, uma importante vantagem dos dados desigualmente espaçados é revelada: o "casamento" de $\cos(2\pi\nu t)$ e $\cos(2\pi\nu' t)$ é exatamente apenas em tempos $t = t_0 + n\Delta t$, então quanto mais os dados se afastarem destes valores de tempo, menor ficará a

fórmula $\sum_n \alpha_n c_n' c_n / \sigma_n^2$. Portanto, um espaçamento irregular assegura que um *alias* refletido vai ser mais fraco que o pico original.

Dos 3 tipos de problemas geradores de *aliases* esboçados, o primeiro não chega a ser um problema real e o terceiro não é importante (já que os dados são geralmente desigualmente espaçados), mas o segundo problema necessita de mais atenção. Faremos a análise proposta pelo método num sinal puramente senoidal, da mesma forma que os dados foram tratados e obtemos um padrão de *aliasing* que seria produzido por um sinal senoidal se esse estivesse presente nos dados. Consideraremos o caso geral de dados que são:

- i) desigualmente espaçados- para tempos t_n ;
- ii) ponderados- por um fator $1 / \sigma_n^2$ para cada dado;
- iii) integrados- sobre um tempo de integração τ_n que pode ser diferente para cada medida (τ_n finito implica atenuação de variações de altas frequências presentes na curva de intensidade de luz).

Para uma maior comodidade, utilizaremos valores iguais de σ_n^2 e τ_n para todas as observações.

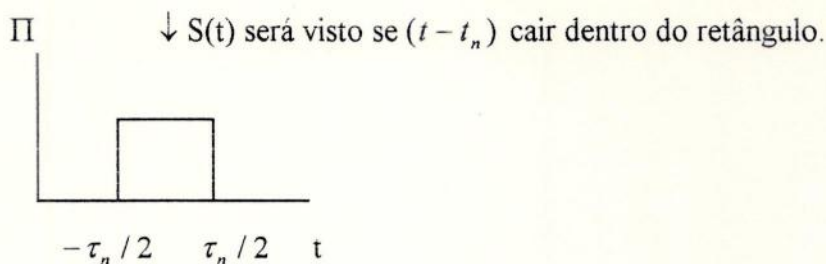
Se a função fonte contínua é $S(t)$ (isto é, o sinal que sai da fonte), então o sinal que iremos realmente medir na nossa n -ésima medida é dado por:

$$f(t_n) = (1 / \tau_n) [S(t) \otimes \Pi((t - t_n) / \tau_n)]$$

onde $S(t)$ está em convolução com a função:

$$\Pi((t - t_n) / \tau_n) = 1, \text{ para } -\tau_n / 2 < t - t_n < \tau_n / 2; \text{ e } 0, \text{ para } -\tau_n / 2 > t - t_n > \tau_n / 2$$

Na verdade, a função Π nos diz se $S(t)$ é captado ou não, isto é, se ele "cai" na janela delimitada pelo intervalo. Se S é medido para um tempo $(t - t_n)$ dentro dos limites do tempo de integração τ_n , isto é, $-\tau_n / 2 < t - t_n < \tau_n / 2$, ele será captado ($\Pi = 1$). Se, por outro lado, S cair fora do intervalo arbitrado pelo tempo de integração, isto é, $t - t_n < -\tau_n / 2$ e $t - t_n > \tau_n / 2$, ele não será captado ($\Pi = 0$).



Considere agora a resposta do sistema para um sinal puramente monocromático de amplitude unitária e fase arbitrária:

$$S_G(t) = \cos(2\pi\nu t + \phi)$$

Substituindo esta fórmula em $f(t_n) = (1/\tau_n)[S(t) \otimes \Pi((t-t_n)/\tau_n)]$ e, calculando as amplitudes Fourier (tais com $A(\nu)$ e $B(\nu)$), teremos:

$$A_G(\nu) = \left[\sum_n \cos(2\pi\nu t_n + \phi) c_n \text{senc}(\nu t_n) / \sigma_n^2 \right] / \left[\sum_n c_n^2 / \sigma_n^2 \right]$$

$$B_G(\nu) = \left[\sum_n \cos(2\pi\nu t_n + \phi) s_n \text{senc}(\nu t_n) / \sigma_n^2 \right] / \left[\sum_n s_n^2 / \sigma_n^2 \right]$$

onde $\text{senc}(x) = \text{sen}(\pi x) / \pi x$.

Então, dado nosso padrão amostral (isto é, os valores de t_n , σ_n e τ_n para $n = 1, 2, \dots, N$) podemos gerar uma estimativa espectral $P_G^2(\nu) = A_G^2(\nu) + B_G^2(\nu)$ que seria produzida se um sinal periódico de dada frequência e fase estivesse presente nos dados; isso ajudaria a identificação entre os picos reais e os picos de *alias*.

c) Programas: Ao contrário do método Jurkevich, que utilizou somente um programa, o método Ponman será aplicado através de um conjunto de vários programas, tendo em vista uma melhor eficiência computacional (um único programa para o método possuiria mais variáveis indexadas que o número de variáveis que um programa normalmente comporta). Todos os programas novamente estão em PASCAL e são divididos em 2 grupos: o primeiro grupo, composto de 5 programas, é responsável pela aplicação do método em si (isto é, o cálculo dos pontos do periodograma, a variância de cada ponto e a aplicação, ou não, do termo de variância associado com o cancelamento imperfeito de termos cruzados). O primeiro programa deste grupo, Ponman1.pas, trata do cálculo dos numeradores de $A(\nu)$ e $B(\nu)$, isto é,

de $\sum_n f_n c_n$ e $\sum_n f_n s_n$, respectivamente além de gerar o intervalo de frequências. O segundo programa, Ponman1a.pas, trata do cálculo dos denominadores de $A(\nu)$ e $B(\nu)$, $\sum_n c_n^2$ e $\sum_n s_n^2$, respectivamente. O terceiro programa, Ponman2.pas, calcula os $A(\nu)$ e $B(\nu)$ propriamente ditos (isto é, ele efetua a divisão entre os numeradores de Ponman1.pas e os denominadores de Ponman1a.pas). O quarto programa, Ponman3.pas, calcula $P(\nu) = A^2(\nu) + B^2(\nu)$ e devolve em forma de arquivo, os pontos $P(\nu)$ e ν para serem plotados em gráfico. O quinto programa, Ponman4.pas, calcula $\varepsilon^2(\nu)$, a variância dos pontos de $P(\nu)$, além de verificar a já citada condição:

$$N \gg 8\sigma_f^2 \sum_n (1/\sigma_n^4) / \sum_n (1/\sigma_n^2)$$

e nos oferece a opção de adicionar ou não o termo de variância devida ao cancelamento imperfeito de termos cruzados (a fórmula adotada para $\varepsilon^2(\nu)$ no programa é a sua aproximação, isto é, $8P(\nu) / \sum_n (1/\sigma_n^2)$, com o fim de tornar o programa menos complexo e facilitar os cálculos).

O segundo grupo de programas (em número de 3), é responsável pelo cálculo dos pontos do periodograma de um sinal puramente senoidal. O primeiro, Alias1.pas, calcula (assim como em Ponman1.pas) os numeradores de $A_G(\nu)$ e $B_G(\nu)$ e o intervalo de frequências. O segundo programa, Alias2.pas, calcula os denominadores de $A_G(\nu)$ e $B_G(\nu)$ e o terceiro programa, Alias3.pas, calcula $P_G(\nu) = A_G^2(\nu) + B_G^2(\nu)$, além de armazenar, em arquivo, $P_G(\nu)$ e ν . Assim, calculamos os pontos do periodograma de um sinal puramente senoidal e comparamos seu padrão com o do periodograma da série que estamos analisando.

Ainda existem 2 programas adicionais: um que subtrai a média dos pontos da série (Saimedia.pas) e um que subtrai tendências lentas da série temporal por meio de um filtro de diferenças, além de calcular $1/T$ e a pseudo-frequência de Nyquist (Filtrodif.pas). Estes programas tiveram que ser aplicados às séries devido ao fato de que o método Ponman assume que a média dos dados da série e qualquer tendência lenta foram subtraídas da série

temporal. Assim, de acordo com as características da série temporal estudada, estes programas são aplicados ou não.

4- APLICAÇÕES E DISCUSSÃO DOS RESULTADOS:

Nesta seção, veremos inicialmente as séries temporais utilizadas para a aplicação dos métodos Jurkevich e Ponman. A seguir, faremos a aplicação do método Jurkevich a elas e posteriormente, a aplicação do método Ponman. Fechamos a seção com a discussão dos resultados das aplicações dos métodos.

4.1- As séries temporais.

Neste trabalho utilizaremos 3 séries temporais. Elas são as seguintes:

- Dados sintéticos para VWCEPHEI: A partir de dados observados da binária eclipsante VWCEPHEI, dados sintéticos desigualmente espaçados foram extraídos a partir de uma aproximação dos dados observados feita por Série de Fourier. Ainda foi adicionado um termo de ruído normalmente distribuído. Estes dados sintéticos foram plotados num gráfico que foi impresso no artigo de Jurkevich e seus pontos foram medidos e anotados com a ajuda de um papel milimetrado vegetal. Como estes pontos da série foram extraídos de um gráfico que foi apresentado no artigo, não se pode assegurar que os pontos extraídos tenham fielmente os mesmos valores dos pontos do gráfico. Estes pontos estão plotados na figura 1 e mostram a característica da série temporal: não existem evidências de tendência e seus pontos flutuam em torno de uma média não-nula. No eixo das ordenadas, temos valores diferenciais de magnitude e no eixo das abcissas, valores de tempo em dias.
- Dados sintéticos para uma curva senoidal: Esta série temporal (figura 2), gerada pelo professor José Adolfo Snajdauf de Campos, é uma série sintética monoperiódica, desigualmente espaçada e de características senoidais. Na ordenada, temos valores diferenciais de magnitude e nas abcissas, valores arbitrários de Datas Julianas. Esta série também não aparenta possuir tendência e sua média também não é nula. Seus pontos estão armazenados em

arquivo (quer dizer, não foram extraídos de um gráfico e submetidos a mais uma incerteza como em VWCEPHEI).

- Dados observacionais de HD93044: Esta série (figura 3) é uma série observada para a variável δ Scuti HD93044. Nas ordenadas, o valor de sua magnitude visual e nas abcissas, a Data Juliana (em unidades de 2447900). Esta série, ao contrário das outras, está dividida em blocos. Ao todo esta série possui 3 blocos, onde a separação entre o primeiro e o segundo é de aproximadamente 10 dias e do segundo para o terceiro de, 1 dia. Ela possui uma certa tendência, além de sua média não ser nula. Seus dados estavam tabelados em artigo (Liu Zong-Li, 1993). Como a figura 3 não mostra cada bloco com clareza (esta só foi plotada para dar uma idéia da localização dos blocos), as figuras 4, 5 e 6 nos mostram o gráfico de cada bloco (ou conjunto), para uma melhor visualização.

Nas páginas seguintes aos gráficos de cada série, temos as listagens de seus pontos.

Estas 3 séries temporais nos dão 3 tipos de casos: no primeiro, temos dados sintéticos obtidos a partir de dados observados de VW CEPHEI; no segundo caso, temos dados totalmente sintéticos e, no terceiro, temos dados observados. Existem então duas séries diferentes (observada e sintética) e uma terceira, que é o meio termo entre as duas diferentes (quer dizer, dados sintéticos tomados a partir de dados observados), o que nos vai dar uma boa diversidade no tocante à análise das séries temporais pelos métodos de determinação de periodicidades e frequências.

4.2- Aplicação do Método Jurkevich às séries temporais.

a) Dados sintéticos de VW CEPHEI.

A figura 7 mostra o periodograma obtido pelo Método Jurkevich para os dados sintéticos de VW CEPHEI. Podemos ver nesta figura que temos um intervalo de períodos tentativa compreendido entre 0,20 e 0,35 dia no eixo das abcissas variando em passos de 0,01 (isto é, 0,20, 0,21, etc.). Foram plotadas 5 curvas sendo que em cada uma delas foi usado um número de caixas (ou grupos) diferente. Temos curvas para 2, 4, 5 e 10 caixas, sendo que a de 10 caixas evidencia mais a determinação do período (o mínimo da curva fica mais evidente).

Isto prova que o aumento do número de caixas melhora a eficiência do método. O período encontrado foi de 0,28 dia. A estatística de Bartlett deu 14,040 e a tabela χ^2 assume, para 9 graus de liberdade (os graus de liberdade são dados por $k - 1$ onde k é o número de caixas) e um nível de significância de 5%, o valor 16,919. Logo, a hipótese de que a amostra de variâncias é homogênea é aceita, o que implica $V_m^2 \cong V^2$ (Teorema de Adição de Variâncias). Isso quer dizer que o mínimo em 0,28 dia não possui um amparo estatístico razoável para sua significância (para que o possuísse, a amostra de variâncias deveria ser heterogênea, isto é, o valor da estatística de Bartlett deveria ser maior que o valor χ^2 para 9 graus de liberdade, rejeitando assim a hipótese de homogeneidade; uma amostra heterogênea de variâncias implicaria em $V_m^2 < V^2$ e estamos interessados em V_m^2 baixos, isto é, a baixa dispersão entre os pontos da série que analisamos e a série hipotética cujo período é o período tentativa).

Um novo periodograma é apresentado na figura 8. Aqui, plotamos a curva somente para o caso de 10 caixas, mas aumentamos o número de pontos calculados (o intervalo de períodos tentativa mostra agora passos de 0,001, isto é, 0,200, 0,201, etc.). Nota-se que a estrutura da curva ficou mais ruidosa. Este ruído é gerado pelo próprio algoritmo do programa. Aqui vemos um leve deslocamento do mínimo de 0,28 dia para 0,285 dia. Agora, aumentaremos ainda mais o número de pontos calculados (passos de 0,0001- 0,2001, 0,2002, etc.) e diminuiremos o intervalo de períodos tentativa (agora vai de 0,27 a 0,29). Os resultados destas alterações podem ser vistos na figura 9. Aqui, temos uma estrutura altamente ruidosa onde aparecem mínimos em 0,2768 dia e 0,2824 dia. Novamente, a estatística de Bartlett aceitou a hipótese de homogeneidade de variâncias: 7,908 e 14,233 respectivamente para períodos 0,2768 e 0,2824 dia. As figuras 10 e 11 mostram a estrutura ruidosa de 0,2824 em maiores detalhes, sendo que na figura 11, temos uma espécie de platô na curva entre 0,2824 e 0,2826. Tomamos o ponto médio desta e o período é determinado como 0,2825 dia.

Concluimos que o período encontrado é de 0,2825, embora não haja suporte estatístico para este período.

b) Dados sintéticos senoidais.

Repetindo o procedimento feito para VW CEPHEI, temos na figura 12 o periodograma da série temporal sintética senoidal. Novamente foram plotadas curvas para 2, 4, 5 e 10 grupos e novamente a curva de 10 grupos determinou o período de forma mais evidente, apesar de os casos de 4 e 5 caixas também o fazerem. Foi encontrado num intervalo de períodos tentativa de 10 a 30 e passos de 1 o período 18 (em unidades arbitrárias de Data Juliana, já que se trata de dados totalmente sintéticos). A estatística de Bartlett acusou um valor de 44,989 contra os 16,919 do teste χ^2 para 9 graus de liberdade e nível de significância de 5%. Logo a amostra é heterogênea e $V_m^2 < V^2$. A figura 13 mostra o periodograma com o mesmo intervalo de 10 a 30 em passos de 0,1 (mais pontos calculados). Novamente evidencia-se o ruído mas o mínimo não sai do valor 18 para o período. Na figura 14, temos, para um intervalo entre 17 e 19 de período, o periodograma com passos de 0,01. Aqui, dois mínimos aparecem: um em 17,99 e outro em 18,18 com valores para a estatística de Bartlett de 32,574 e 41,748 respectivamente. Logo, os dois mínimos possuem amparo estatístico e o período será determinado pelo mínimo que possuir o menor valor de V_m^2 , que neste caso é 18,18. A figura 15 nos dá uma visão ampliada destes 2 mínimos.

c) Dados observados de HD 93044.

Assim como em VW CEPHEI e na série sintética senoidal, a figura 16 nos dá o periodograma para a série temporal observada da variável δ Scuti HD 93044. Aqui, mais uma vez a curva de 10 caixas é a que evidencia mais o período embora as restantes também o façam de forma convincente. Foi encontrado um período de 0,09 dia num intervalo de períodos tentativa de 0,04 a 0,13 dia variando em passos de 0,01. O valor para a estatística de Bartlett foi de 18,233 contra os 16,919 do teste χ^2 para 9 graus de liberdade e nível de significância de 5%, dando novamente amparo estatístico para o período determinado. A figura 17 mostra o periodograma com um maior número de pontos calculados (passos de 0,001) para o mesmo intervalo de 0,04 a 0,13 dia. Nota-se que começa a se evidenciar uma certa estrutura regular no periodograma. Esta estrutura torna-se muito mais evidente na figura

18. Aqui, o número de pontos foi aumentado mais ainda (passos de 0,0001) mas o intervalo de períodos tentativa foi diminuído (0,068 a 0,092 dia). A estrutura gerada impressiona pela sua forma regular e "periódica". Os 2 conjuntos de 2 mínimos foram analisados pela estatística de Bartlett. No primeiro conjunto (em períodos de 0,0769 e 0,0775), os valores da estatística de Bartlett foram respectivamente de 10,058 e 15,408 contra os 16,919 do χ^2 . Logo, as amostras são homogêneas. No segundo conjunto de mínimos (em períodos de 0,0833 e 0,084) os valores para Bartlett foram, respectivamente, 20,076 e 6,684. O menor valor de V_m^2 é o que aponta o período 0,084 mas este possui uma amostra de variância homogênea ($6,614 < 16,919$). O mínimo que aponta o período 0,0833 possui uma amostra heterogênea ($20,076 > 16,919$). Portanto, graficamente o período é determinado para um valor 0,084 dia mas o período de 0,0833 dia é o que possui um amparo estatístico e este será considerado o período da série temporal (apesar de a estatística de Bartlett somente verificar se a amostra de variâncias é heterogênea ou não).

Entretanto, uma dúvida paira no ar: por que o periodograma de HD 93044 apresentou uma estrutura tão regular? Sabemos que a série de HD 93044 possui duas lacunas: a primeira separa o primeiro conjunto de observações do segundo por uma extensão de 10 dias aproximadamente e a segunda lacuna separa o segundo conjunto de observações do terceiro por um dia. A série foi então dividida em 3 partes: uma com o primeiro conjunto de observações, outra com o segundo conjunto e a terceira, com o terceiro conjunto. Além disso, foi gerada uma série com o segundo e terceiro conjuntos unidos. Para cada nova "série" foi plotado um periodograma. A figura 19 mostra o periodograma do primeiro conjunto de observações com intervalo de períodos tentativa de 0,04 a 0,13 em passos de 0,001. A figura 20a nos dá o mesmo periodograma agora com passos de 0,0001 e intervalo de 0,068 a 0,092. Note que existe uma estrutura ruidosa que não se parece com o periodograma da figura 18, com o mesmo passo e intervalo de períodos. Isso também ocorre nas figuras 20b e 20c onde são plotados os periodogramas do segundo e terceiro conjunto de observações, respectivamente. Vamos agora à figura 21, que é o periodograma da série que comporta

juntos o segundo e terceiro conjunto de observações. Aqui, para um mesmo passo e intervalo dos casos anteriores, vemos que o periodograma repete novamente um certo padrão, não igual ao da figura 18 mas semelhante. Uma característica comum às duas séries destes periodogramas semelhantes é que elas possuem lacunas. Quando foi plotado um periodograma de dados sem lacunas, a estrutura desapareceu. Logo, a presença de uma "regularidade" nos periodogramas talvez possa ser atribuída à presença de lacunas nos dados.

4.3- Aplicação do Método Ponman às séries temporais.

a) Dados sintéticos de VW CEPHEI.

Antes de mais nada, torna-se necessário analisar a série visualmente. Isso é importante devido ao fato de que toda a série temporal que se submeter à análise deste método deve possuir média nula e qualquer tendência lenta subtraída. Uma análise de VW CEPHEI nos mostra que, aparentemente não existem tendências lentas na série, mas sua média não é nula. Subtraímos a média dos dados com a ajuda do programa Saimedia.pas e plotamos a série temporal novamente (figura 22). Agora esta série temporal está apta a ser analisada pelo Método Ponman. A figura 23 mostra o periodograma gerado pelo Método Ponman para a série sintética de VW CEPHEI. Aqui, o eixo das abcissas está em unidades de frequência e seu intervalo obedece o seguinte critério: seu limite inferior é $1/T$, onde T é o tamanho total da série no eixo das abcissas; e seu limite superior é a pseudo-frequência Nyquist $\nu_{Ny} = (2 \langle \Delta t \rangle)^{-1}$ (este critério é seguido somente com o objetivo de se obter uma boa resolução para o periodograma; máximos podem estar fora deste intervalo). O intervalo está então entre 0,848 e 19 (um pouco acima da pseudo-frequência de Nyquist, que é 18,644). O periodograma nos mostra o seguinte: um grande máximo em 7,131 ciclos/dia (aproximadamente um período de 0,14 dia) e 3 picos satélites. O primeiro (em aproximadamente 3,4 ciclos/dia) corresponde a um período em torno de 0,28, 0,29 dia, sendo, portanto, o período determinado pelo Método Jurkevich. Como o período principal está na

casa de 0,14 dia (7,131 ciclos/dia), podemos dizer que estes dois picos são harmônicos. O segundo pico satélite está em, aproximadamente, 9,1 ciclos/dia e não parece ter nenhum significado mais evidente. O terceiro pico está nas imediações de 18,64 ciclos/dia (18,2 a 18,5), a pseudo-frequência Nyquist e pode estar sendo provocado por ela. Outra análise que podemos fazer é plotar um periodograma aplicando o Método Ponman não aos pontos da série temporal, mas à uma função janela senoidal (vide item de aproximações e erros do Método Ponman) cuja frequência é a determinada pelo periodograma do Método Ponman aplicado aos pontos da série (neste caso, 7,131 ciclos/dia). Assim, o periodograma gerará uma curva que ocorreria se a série temporal que analisamos fosse senoidal. O periodograma de uma função senoidal funciona como uma espécie de "impressão digital" e uma comparação entre o padrão do periodograma da série observada e o padrão do periodograma da função senoidal nos diz o quanto a série temporal pode possuir ou não características senoidais. A figura 24 é um periodograma para uma função senoidal com frequência 7,131 ciclos/dia. Outros parâmetros utilizados foram: tempo de integração τ de 0,00011 dia e fase nula (vide fórmulas de A_G e B_G no item de aproximações e erros de Ponman); os desvios-padrão para cada ponto foram arbitrados como iguais, eliminando os termos $\sum_n 1/\sigma_n^2$ nas fórmulas de A_G e B_G). Notamos, nesta figura, várias características importantes: além do pico principal, obviamente situado em 7,131 ciclos/dia, vemos repetido o pico secundário de 9,1 ciclos/dia inclusive com o mesmo formato; só está um pouco menor. Além disso, o pico em 3,4 ciclos/dia desapareceu (a série senoidal só possui uma frequência, não possui harmônicos) e o pico que era candidato à pseudo-frequência Nyquist desapareceu também. Dentre essas características, a mais importante sem dúvida é a repetição do pico secundário em 9,1 ciclos/dia. Este pico (e mais algumas pequenas outras semelhanças entre os dois periodogramas) indicam uma pequena característica senoidal à série sintética de VW CEPHEI.

A variância do pico em 7,131 ciclos/dia (onde não foi considerado o termo de variância devido ao cancelamento imperfeito de termos cruzados) foi calculada pela fórmula

aproximada de $8P(\nu) / (\sum_n 1/\sigma_n^2)$ (onde σ_n^2 foi arbitrado como 0,01 para todos os dados) tendo o valor de 0,000035 sendo, portanto, muito pequena em comparação com o valor do máximo que é de 0,01884 e dando um amparo estatístico para o máximo.

Temos novamente outra dúvida: será que no Método Jurkevich, o período 0,14 dia tinha algum mínimo? Lembremos que o intervalo utilizado na análise foi de 0,20 a 0,35 dia. Para esclarecer esta dúvida, plotamos novamente o periodograma Jurkevich para VW CEPHEI agora para um intervalo de 0,10 a 0,35 dia, como mostra a figura 25. Nota-se que em 0,14 dia também aparece um mínimo, só que com valor de V_m^2 um pouco maior que o valor em 0,28 dia (0,1221179 para 0,14 dia contra 0,1184907 para 0,28 dia). A estatística de Bartlett acusou 7,208 para 0,14 dia contra os 16,919 do teste χ^2 para 9 graus de liberdade e nível de significância de 5%, sendo que este mínimo, assim como o de 0,28 dia não possui amparo estatístico. Entretanto, a análise da figura 25a, que é um periodograma Jurkevich para um intervalo de 0,1370 a 0,1440 com passos de 0,0001, nos mostra um mínimo em 0,1407 e com um valor de 0,116888! Além disso, a estatística de Bartlett acusou 33,267. Concluímos então que o período determinado por Jurkevich não é 0,2825 dia como se pensava e sim 0,1407 dia.

b) Dados sintéticos senoidais.

Uma inspeção visual da série sintética temporal com características senoidais mostra que não parece existir tendência lenta; entretanto, a média não é nula. Portanto, faz-se novamente a eliminação da média. O resultado aparece na figura 26. Aplicamos o Método Ponman e o seu periodograma aparece na figura 27, com intervalo de frequências entre 0,021 e 1 (1/T e pseudo-frequência Nyquist, aproximadamente). Aqui, vemos um máximo evidente em 0,053 ciclos/dia, o que corresponde a um período de 18,87 dias estando em relativa concordância com o período encontrado por Jurkevich, que foi de 18,18. Além deste máximo, nota-se uma estrutura formada por uma cadeia de picos secundários. Tomamos, então, o periodograma da função janela senoidal para uma frequência 0,053 e os mesmos parâmetros usados em VW CEPHEI: fase nula e tempo de integração de 0,00011 dia (figura

28). O resultado nos mostra boa concordância entre o padrão de picos secundários do periodograma dos pontos da série e os do periodograma da função senoidal (muitos deles se repetem). Isso mostra que a verificação da série temporal possuir características senoidais ou não por comparação entre o padrão dos periodogramas é válida, já que sabemos de antemão que a série analisada é senoidal e que os periodogramas concordam seus padrões em grande parte. A única discórdia foi um pico mais evidente em 0,95 ciclos/dia no periodograma da função senoidal. A variância para o pico principal sem o termo atribuído ao cancelamento imperfeito de termos cruzados foi de 0,000000156 para $\sigma_n^2 = 0,001$, novamente sendo muito baixa em comparação com o valor do máximo que é 0,001737.

c) Dados observados de HD 93044.

A série temporal de HD 93044 mostra, ao contrário das outras, uma certa tendência, além de também estar com média não-nula. Aplicando o programa de redução da média Saimedia.pas e, posteriormente, um filtro de diferenças através do programa Filtrodif.pas, a série ficou com média nula e sem tendência como mostra a figura 29. Seu periodograma aparece na figura 30. Aqui, o intervalo tem como valor mínimo $1/T=0,019$ e valor máximo 23 (este é um caso onde o período está fora do intervalo delimitado pela pseudo-frequência Nyquist e $1/T$ mostrando que este intervalo é usado somente para garantir uma mínima resolução para o periodograma). A frequência determinada foi de 13,002 ciclos/dia, correspondendo a um período de 0,0769 dia que está próximo ao que foi encontrado por Jurkevich (0,0833 dia). O periodograma da função senoidal com frequência 13,002 ciclos/dia, fase nula e tempo de integração 0,00011 dia está na figura 31. Note novamente grande semelhança entre os padrões dos periodogramas, com a exceção de que a extremidade do periodograma dos pontos observados em torno de 23 ciclos/dia possui certos picos secundários. Isso novamente indica que a série observada possui características senoidais. Outra coisa que devemos ressaltar no periodograma é a sua estrutura regular, assim como ocorreu no periodograma gerado pelo Método Jurkevich.

A variância para o máximo em 13,002 ciclos/dia foi de 0,00000000081, sendo novamente muito pequena em comparação com o valor do máximo que é 0,000048561. Observação: Aqui foi adicionado o termo de variância relacionado ao cancelamento imperfeito de termos cruzados, já que a série não é semi-regular (isto é, contém lacunas).

4.4- Discussão dos resultados.

Vista a aplicação dos métodos às séries temporais, podemos dizer que os períodos determinados pelo Método Jurkevich para a série sintética senoidal e para HD 93044 (18,18 dias e 0,0833 dia, respectivamente) estão em boa concordância com os períodos determinados pelo Método Ponman para as mesmas séries (18,87 dias e 0,0769 dia, respectivamente). Os períodos apresentados em Jurkevich para estas duas séries contaram também com um amparo estatístico à significância de seus mínimos de V_m^2 , isto é, a estatística de Bartlett confirmou a hipótese de amostra heterogênea de variâncias, afirmando estatisticamente que V_m^2 era um valor pequeno. É certo que não podemos nos apoiar muito neste teste (ele só confirma se a amostra de V_i^2 é heterogênea ou não), mas como o teste-F é utilizado com restrições (Jurkevich, 1971) e a estatística de Bartlett tem como única restrição a presença de pelo menos dois dados em cada caixa, esta última foi utilizada como indicador da "pequenez" de V_m^2 .

Com relação à VW CEPHEI, os dois métodos apontaram valores extremos de suas "funções sensíveis à períodos" para mesmos períodos e frequências. Entretanto, a intensidade desses extremos foi diferente em cada método. Em Jurkevich, foi apontado um mínimo em 0,28 dia e outro em 0,14 dia. Numa primeira aproximação, o método determina o período de 0,28 como sendo o da série; entretanto, para uma segunda aproximação (aumentando o número de passos), vemos que o período é 0,1407 dia e não 0,28 dia, embora a diferença entre os seus mínimos seja muito pequena. No Método Ponman, os máximos do periodograma apontaram frequências cujos períodos também estão na casa de 0,28 dia e 0,14 dia; entretanto, o máximo para 0,14 dia é muito mais evidente que o de 0,28 dia, ao contrário do que ocorre no Método Jurkevich onde a diferença entre os mínimos dos dois períodos é muito pequena.

Os métodos demonstraram eficiência ao denunciarem igualmente em seus periodogramas o padrão de regularidade para a análise de HD 93044, isto é, os dois periodogramas indicaram as características daquela série com muita clareza.

5- CONCLUSÕES:

Com relação à aplicabilidade dos métodos às séries, podemos dizer que, à exceção do problema ocorrido para VW CEPHEI, ambos os métodos são eficientes para a determinação de periodicidades das séries temporais aqui apresentadas. Apenas cabe aqui ressaltar que em termos de fornecimento de informações, o Método Ponman é mais completo que o Jurkevich, já que o último apenas indica a periodicidade e o primeiro, além de indicar a periodicidade, pode nos dizer se a estrutura da série possui características senoidais ou não, pode indicar padrões de *aliases* e também a pseudo-frequência Nyquist, que pode ser considerada como uma espécie de "essência" da série temporal analisada. Pode-se dizer que foi alcançado o objetivo de se verificar a eficiência dos métodos para cada série analisada.

Essas aplicações apontam também para o aviso dado por Kurtz (1983) quanto aos resultados de cálculos de frequências em séries medidas sem um adequado dimensionamento em amostragem que respeita o limite informacional exigido pela pseudo-frequência de Nyquist.

6- AGRADECIMENTOS:

Expresso meus mais sinceros e profundos agradecimentos às pessoas aqui citadas:

- ☛ Ao Professor Jorge de Albuquerque Vieira, pela orientação sensacional dada a este trabalho e por sempre estar à disposição para esclarecer minhas dúvidas;
- ☛ À Professora Encarnación Amelia Martinez Gonzalez, que me iniciou nesta área de pesquisa e sempre me ajudou a esclarecer as dúvidas em Estatística (além de ser uma excelente revisora de textos);

☉ Ao Professor José Adolfo Snajdauf de Campos, por ter me cedido a série temporal sintética senoidal e por ter me prestado esclarecimentos muito úteis referentes aos métodos *String-Lenght*.

7- BIBLIOGRAFIA:

Deeming, T., J., (1975) Fourier Analysis With Unequally Spaced Data, *Ap. Space Sci.* **36**, p. 137

Dixon, W. D. , Massey Jr., F. J., (1983) Introduction To Statistical Analysis. (Fourth Edition), (*Mc-Graw Hill, New York*), p. 358

Dworetzky, M. M. (1983), A Period Finding Method For Sparse Randomly Spaced Observations *Mon. Not. R. Astr. Soc.*, **203**, p. 917

Fullerton, A. W. (1987), Searching For Periodicity In Astronomical Data, *Asteroids II*, Ed. R. Binzel, p. 201

Hamilton, W. C. (1964), Statistics In Physical Science (*The Ronald Press Co. New York, N.Y.*)

Hoyng, P., (1976), *Astr. Astrophys.* **47**, p. 449

Jurkevich, I. (1971), A Method Of Computing Periods Of Cyclic Phenomena, *Astrophys and Space Sci.* **13**, p. 154.

Kurtz, D., W., (1983), The High Frequency Limit To Fourier Analysis. A Reminder Of The Nyquist Frequency, *Commision 27 of the I. A. U. Information Bulletin on Variable Stars No. 2285*, p.1

Lafler, J., Kinman, T., D., (1965), An RR Lyrae Star Survey With The Lick 20- Inch Astrograph By Electronic Computer , *Astrophys. Suppl. Ser.* **11**, p. 216

Liu Zong-Li, (1993), The Period Analysis of HD 93044 and Its Amplitude Variations, *Astr. and Astrophys.*, **274**, p. 220

Pereira, B., B., (1984), Séries Temporais Multivariadas, *Sexto Simpósio Nacional de Probabilidade e Estatística, Instituto de Matemática, UFRJ.*

Pereira, B., B, Pais, M., B., Z., Sales, P., R., H., (1986), Análise Espectral de Séries Temporais , *Eletróbrás, Arte Final.*

Ponman, T., (1981), The Analysis of Periodicities in Irregularly Sampled Data, *Mon. Not. R. Astr. Soc* **196**, p. 583

Renson, P. (1978), *Astr. Ap.*, **63**, p. 125

Scargle, J. D., (1982), Studies In Astronomical Time Series Analysis II. Statistical Aspects Of Spectral Analysis Of Unevenly Spaced Data, *The Astrophysical Journal*, **263**, p. 835

Stellingwerf, R., F., (1978), Period Determination Using Phase Dispersion Minimization, *The Astrophysical Journal*, **224**, No. 3, p. 953

Vanicek, P., (1969), Approximate Spectral Analysis By Least-Squares Fit (Successive Spectral Analysis), *Astr. And Space Sci.* **4**, p. 387

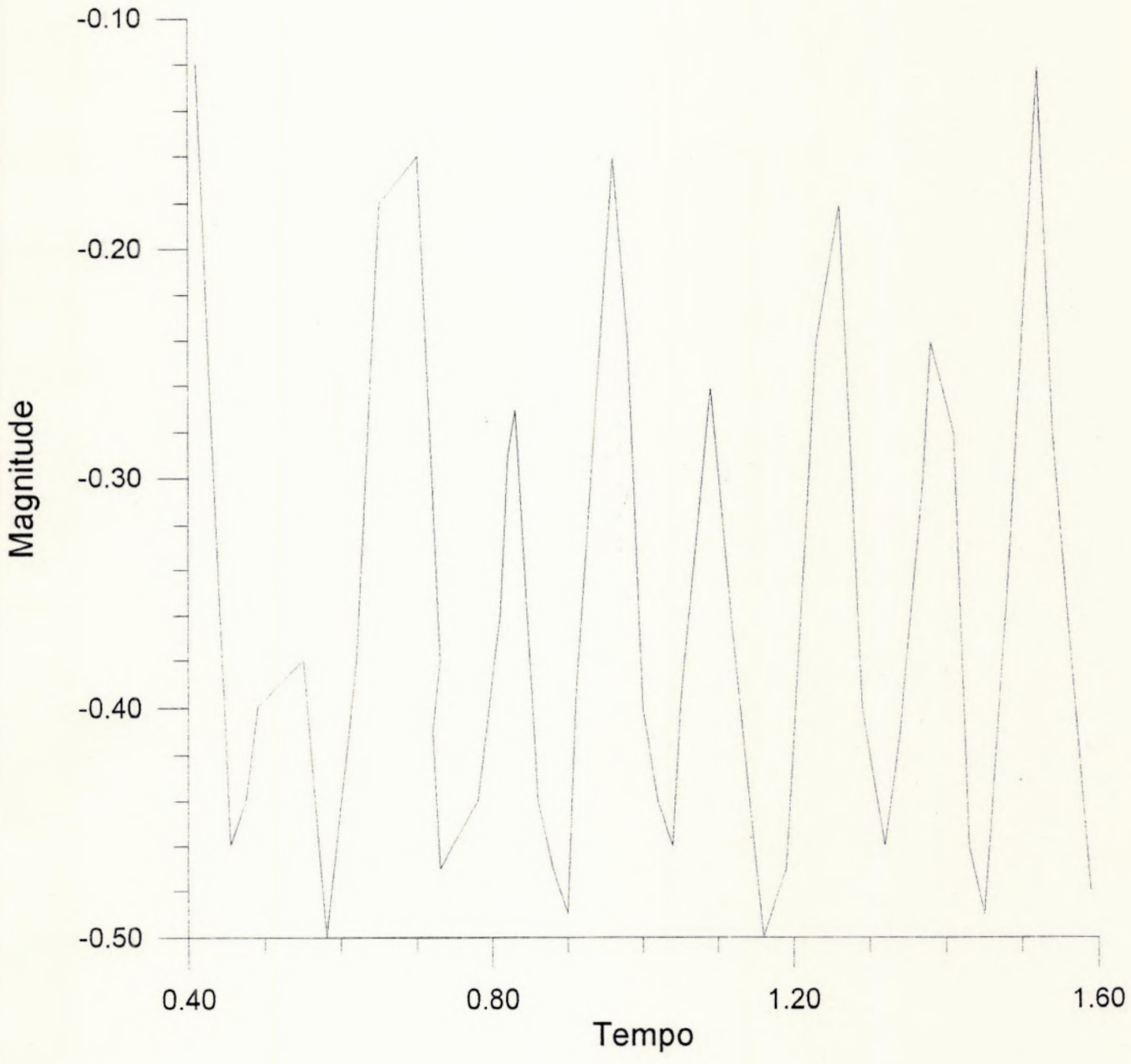


Fig. 1- Série temporal para VW CEPHEI

0.41	-0.12
0.43	-0.28
0.455	-0.46
0.475	-0.44
0.49	-0.40
0.55	-0.38
0.58	-0.50
0.62	-0.38
0.65	-0.18
0.70	-0.16
0.71	-0.38
0.72	-0.41
0.73	-0.47
0.78	-0.44
0.81	-0.36
0.82	-0.29
0.83	-0.27
0.86	-0.44
0.88	-0.47
0.90	-0.49
0.91	-0.40
0.96	-0.16
0.98	-0.24
1.00	-0.40
1.02	-0.44
1.04	-0.46
1.05	-0.40
1.09	-0.26
1.12	-0.37
1.16	-0.50
1.19	-0.47
1.23	-0.24
1.26	-0.18
1.29	-0.40
1.32	-0.46
1.34	-0.41
1.37	-0.30
1.38	-0.24
1.41	-0.28
1.43	-0.46
1.45	-0.49
1.48	-0.35
1.52	-0.12
1.54	-0.28
1.59	-0.48
'tpo'	'magn'

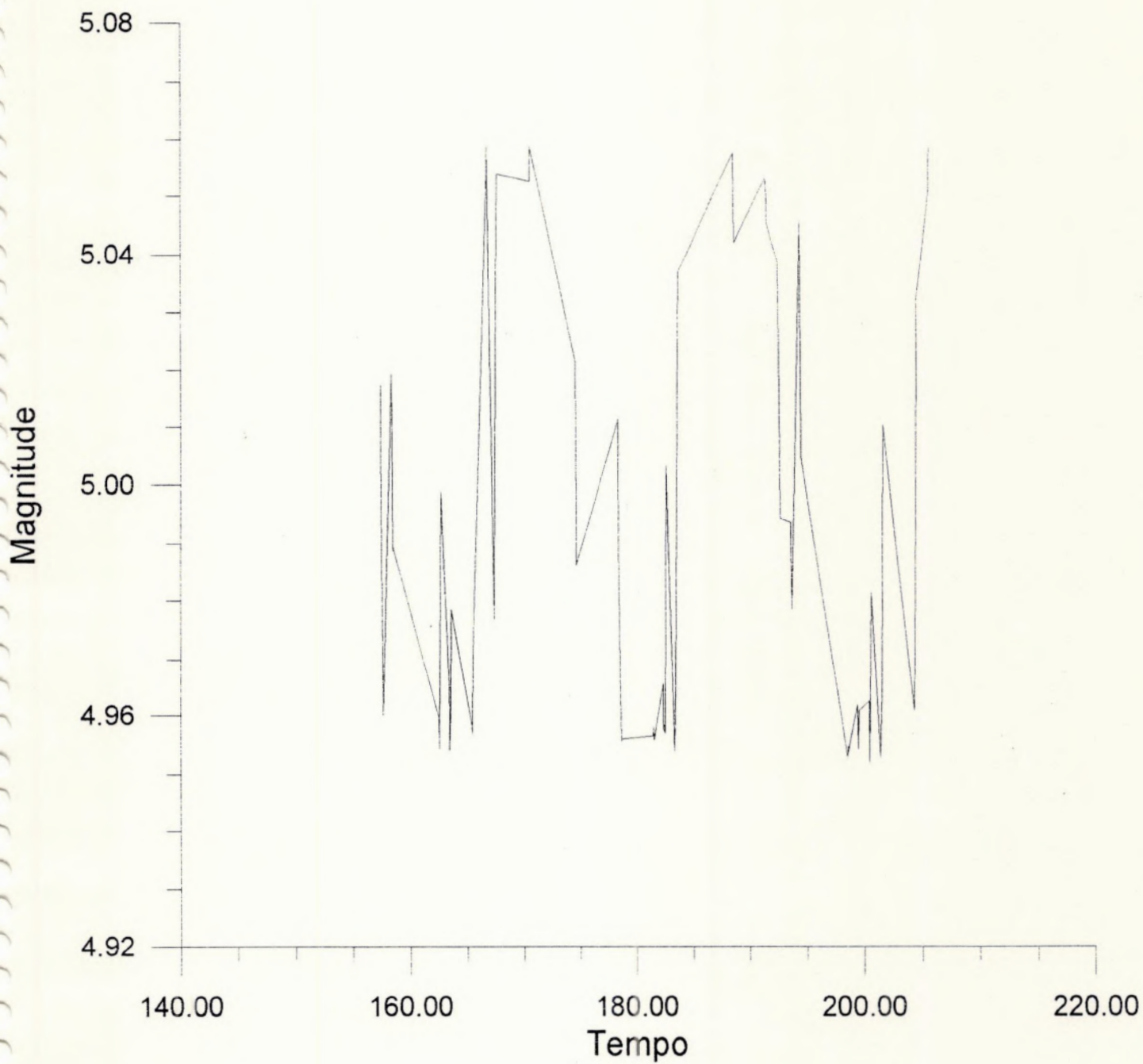


Fig. 2- Série Temporal sintética Senoidal

157.3667 5.017228
157.4055 5.00662
157.6033 4.960094
158.3122 5.019298
158.4299 4.988716
158.4393 4.989457
162.4425 4.959573
162.5044 4.95448
162.6374 4.998987
163.3173 4.965159
163.3923 4.958974
163.3931 4.954124
163.4182 4.957288
163.4804 4.960204
163.5263 4.978535
165.3438 4.95823
165.3705 4.957003
165.4137 4.97103
166.6045 5.051105
166.6287 5.054085
166.6302 5.058848
167.3204 4.976755
167.3406 4.978595
167.3899 5.003653
167.4732 5.034395
167.5607 5.053833
170.4231 5.052493
170.461 5.052595
170.4668 5.058774
174.4745 5.021669
174.5735 4.993795
174.6062 4.986198
178.3101 5.011272
178.3207 5.004587
178.3777 4.990666
178.496 4.967843
178.6116 4.955763
178.6129 4.956079
181.4443 4.956628
181.4665 4.958105
181.5104 4.955774
182.307 4.965571
182.3265 4.959378
182.3475 4.957758
182.4746 4.957043
182.5864 5.002727
182.5947 5.003316
183.3476 4.953881
183.3745 4.959457
183.4592 4.966236
183.5386 4.994522
183.5759 5.0147

183.6356 5.036983
188.5108 5.057566
188.5449 5.048802
188.6028 5.042103
191.3035 5.053077
191.434 5.049934
191.4513 5.045521
192.4324 5.038677
192.4959 5.023772
192.6281 4.99445
193.5806 4.993749
193.5922 4.990017
193.633 4.978455
194.3093 5.045587
194.3447 5.034762
194.4655 5.004902
198.5305 4.952935
198.5735 4.954873
198.579 4.95384
199.4043 4.961917
199.4247 4.957287
199.4434 4.959136
199.4752 4.954373
199.5208 4.960437
199.5388 4.96079
200.3936 4.962605
200.4483 4.952221
200.6024 4.981392
201.3978 4.953115
201.5059 4.963963
201.606 5.0103
204.3097 4.960819
204.4414 4.995499
204.5089 5.03241
205.5619 5.051132
205.5688 5.056212
205.6157 5.058701
'tempo' 'magn'

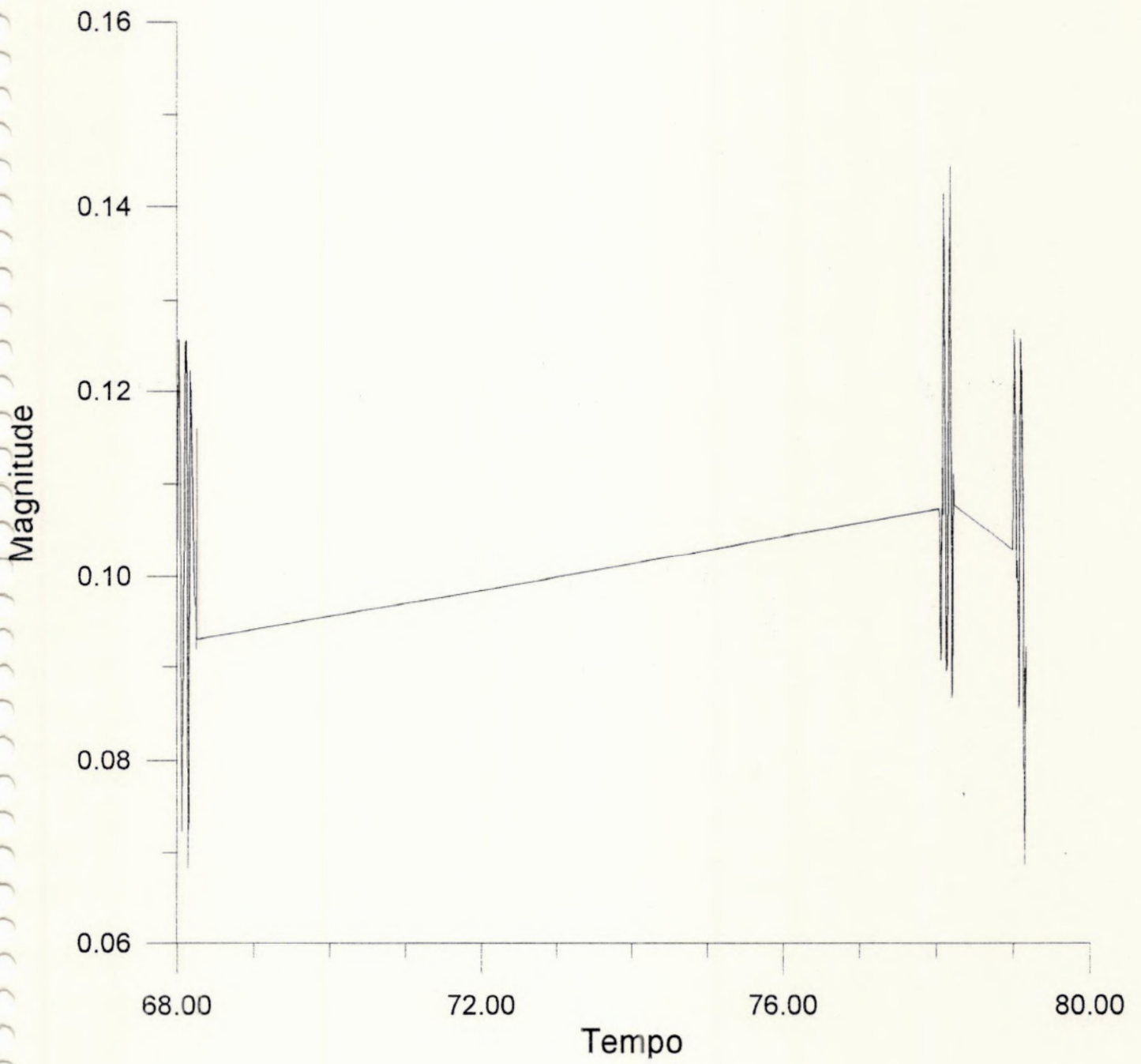


Fig. 3- Série Temporal de HD93044 (dividida em "lacunas")

(2448300 +)

2448368.00482	.1039
2448368.01132	.1158
2448368.01697	.1258
2448368.02516	.1182
2448368.03042	.1257
2448368.03704	.1176
2448368.04312	.1114
2448368.04956	.0934
2448368.05515	.0843
2448368.06383	.0723
2448368.06903	.0823
2448368.07471	.0829
2448368.08043	.0942
2448368.08621	.0963
2448368.09073	.1020
2448368.09564	.1035
2448368.10042	.1231
2448368.10547	.1251
2448368.11128	.1256
2448368.11622	.1221
2448368.12179	.1257
2448368.12698	.1211
2448368.13296	.1032
2448368.13822	.0900
2448368.14321	.0683
2448368.14818	.0766
2448368.15310	.0732
2448368.15775	.0871
2448368.16255	.1040
2448368.16783	.1223
2448368.17354	.1160
2448368.17888	.1210
2448368.18450	.1194
2448368.22528	.0987
2448368.23160	.0970
2448368.23699	.0967
2448368.24325	.0978
2448368.24851	.0920
2448368.25341	.1159
2448368.25879	.0930
2448378.03850	.1074
2448378.04414	.1069
2448378.04864	.1060
2448378.05250	.0945
2448378.05616	.0941
2448378.06029	.0941
2448378.06427	.0908
2448378.06759	.0924
2448378.07198	.1038
2448378.07625	.1030
2448378.08026	.1030
2448378.08381	.1119

2448378.08776	.1068
2448378.09203	.1266
2448378.09715	.1319
2448378.10213	.1417
2448378.10601	.1387
2448378.11037	.1354
2448378.11405	.1253
2448378.11753	.1259
2448378.12087	.1244
2448378.12434	.1105
2448378.12916	.1030
2448378.13287	.0918
2448378.13694	.0900
2448378.14056	.0896
2448378.15320	.0904
2448378.15790	.1013
2448378.16300	.1124
2448378.16759	.1209
2448378.17168	.1277
2448378.17593	.1358
2448378.17982	.1339
2448378.18364	.1334
2448378.18828	.1448
2448378.19196	.1310
2448378.19566	.1210
2448378.20027	.1171
2448378.20417	.1101
2448378.20872	.0952
2448378.21293	.0867
2448378.21745	.0876
2448378.22146	.0949
2448378.22521	.1056
2448378.22909	.1079
2448378.23292	.1112
2448378.23691	.1078
2448378.99977	.1031
2448379.00613	.1096
2448379.01163	.1160
2448379.01698	.1270
2448379.02152	.1243
2448379.02748	.1205
2448379.03602	.1148
2448379.04115	.1042
2448379.04578	.1000
2448379.05063	.1011
2448379.05576	.0996
2448379.06035	.1018
2448379.07512	.0889
2448379.07949	.0856
2448379.08256	.0863
2448379.08683	.1022
2448379.09134	.1140

2448379.09506	.1223
2448379.09874	.1260
2448379.10241	.1257
2448379.10724	.1204
2448379.11102	.1176
2448379.11488	.1224
2448379.11895	.1140
2448379.12296	.1170
2448379.12626	.1140
2448379.12951	.1077
2448379.13315	.1031
2448379.13635	.0965
2448379.14015	.0840
2448379.14334	.0738
2448379.14718	.0705
2448379.15015	.0686
2448379.15345	.0795
2448379.15704	.0790
2448379.16056	.0900
2448379.16394	.0840
2448379.16716	.0923
'tempo'	'magn'

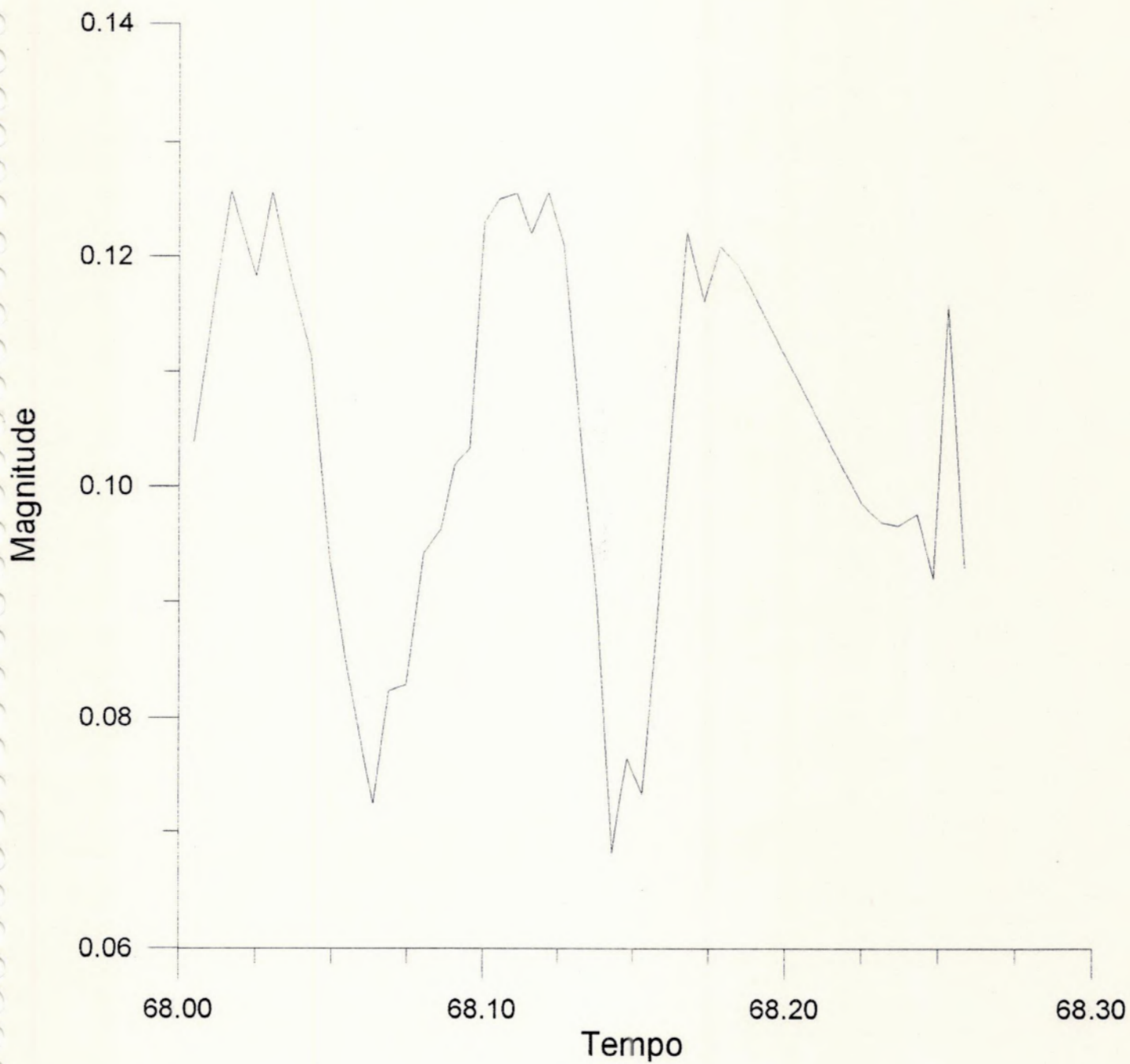


Fig. 4- Primeiro conjunto de hd93044
(2448300 +)

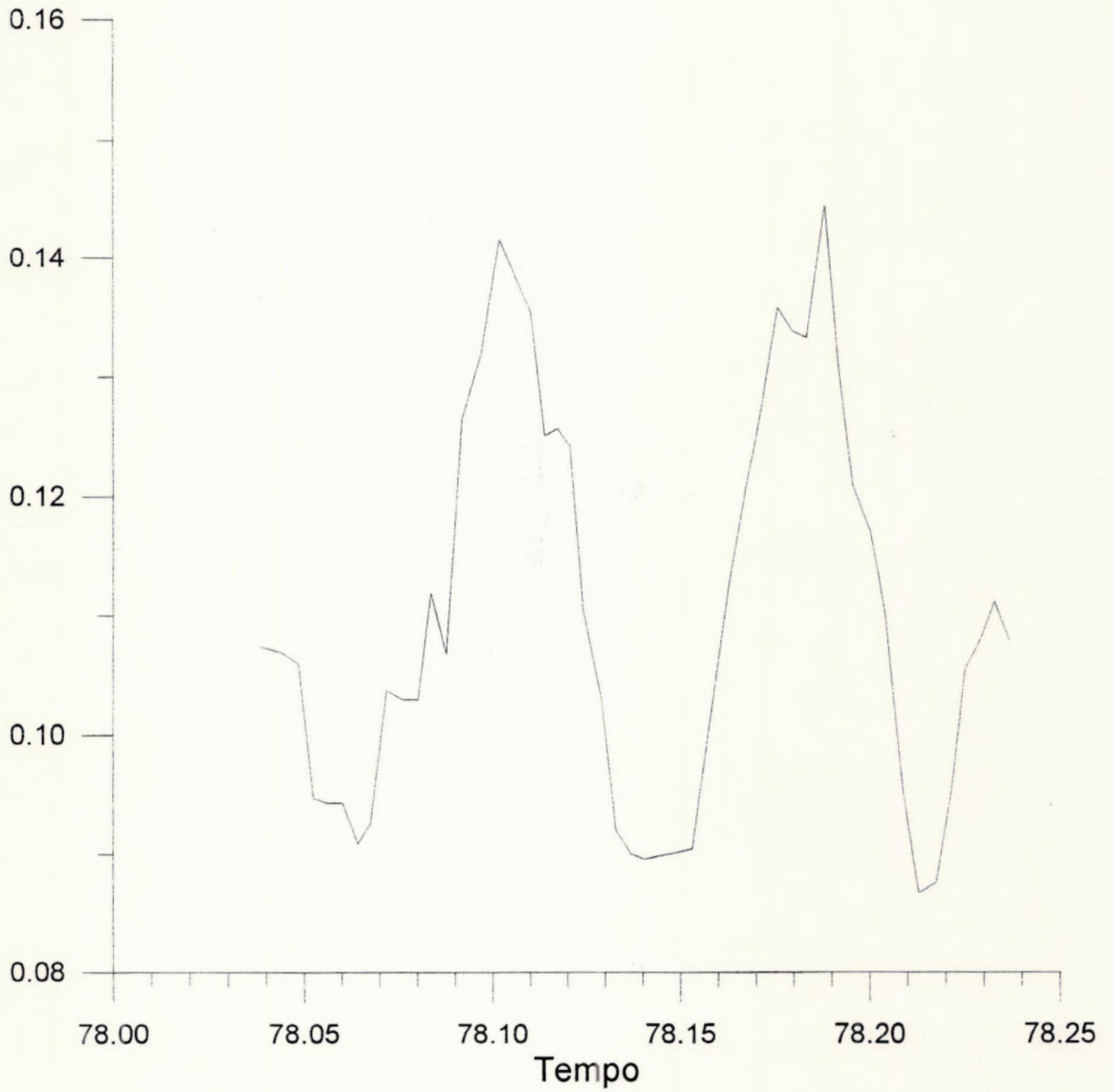


Fig. 5- Segundo conjunto de HD93044

(2448300 +)

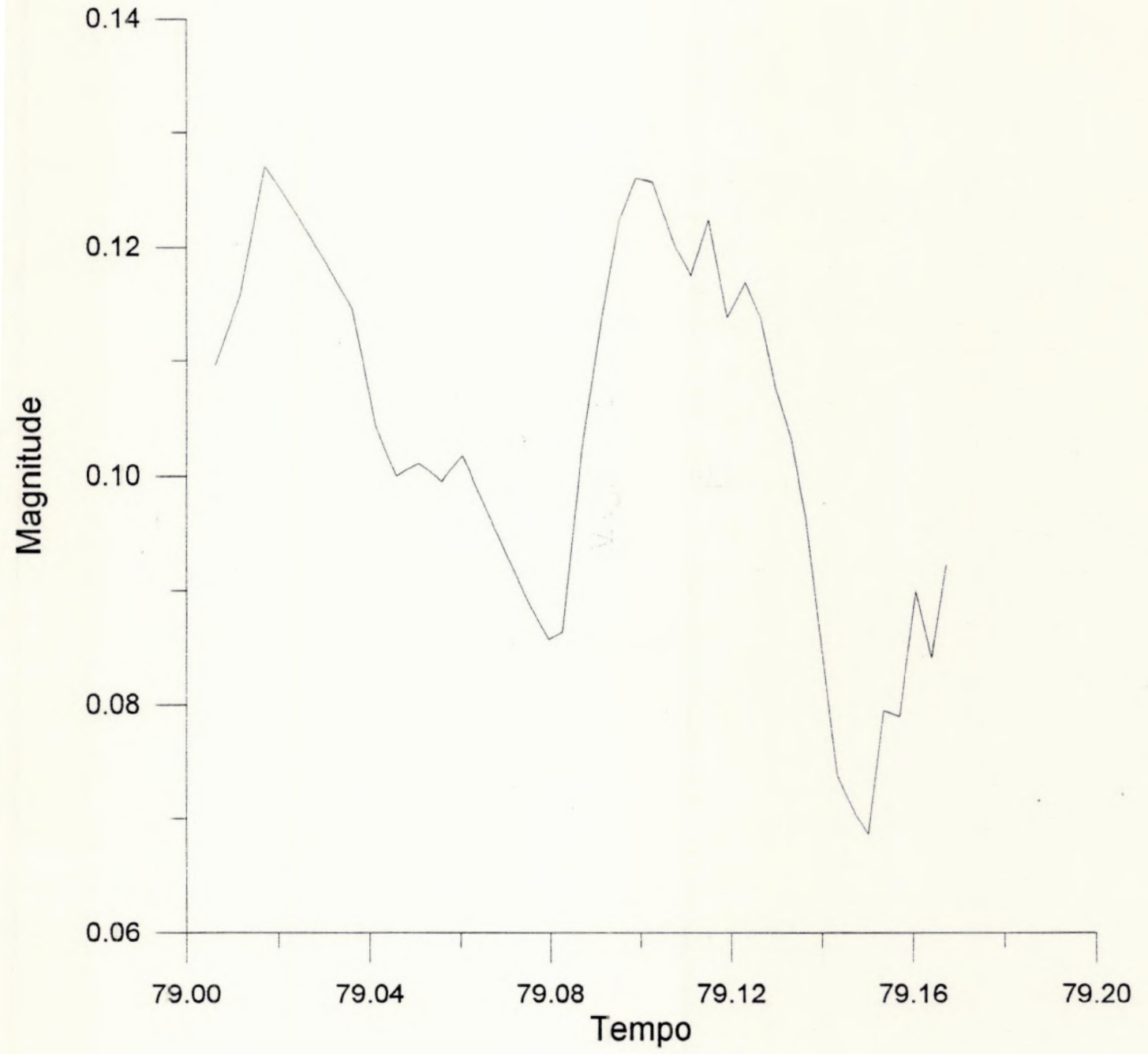


Fig. 6- Terceiro conjunto de HD93044

(2448300 +)

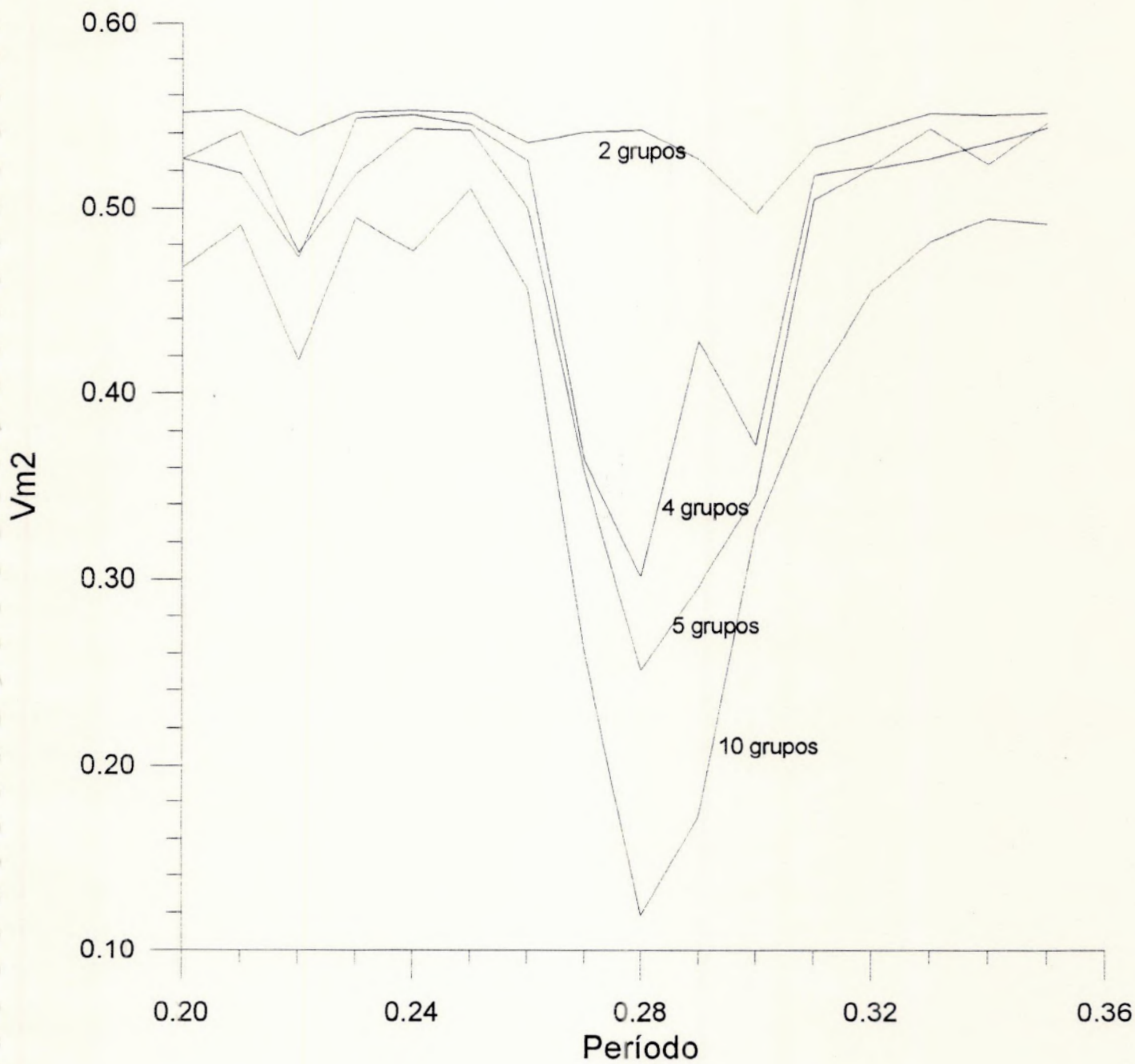


Fig. 7- Periodograma Jurkevich para 2, 4, 5 e 10 grupos

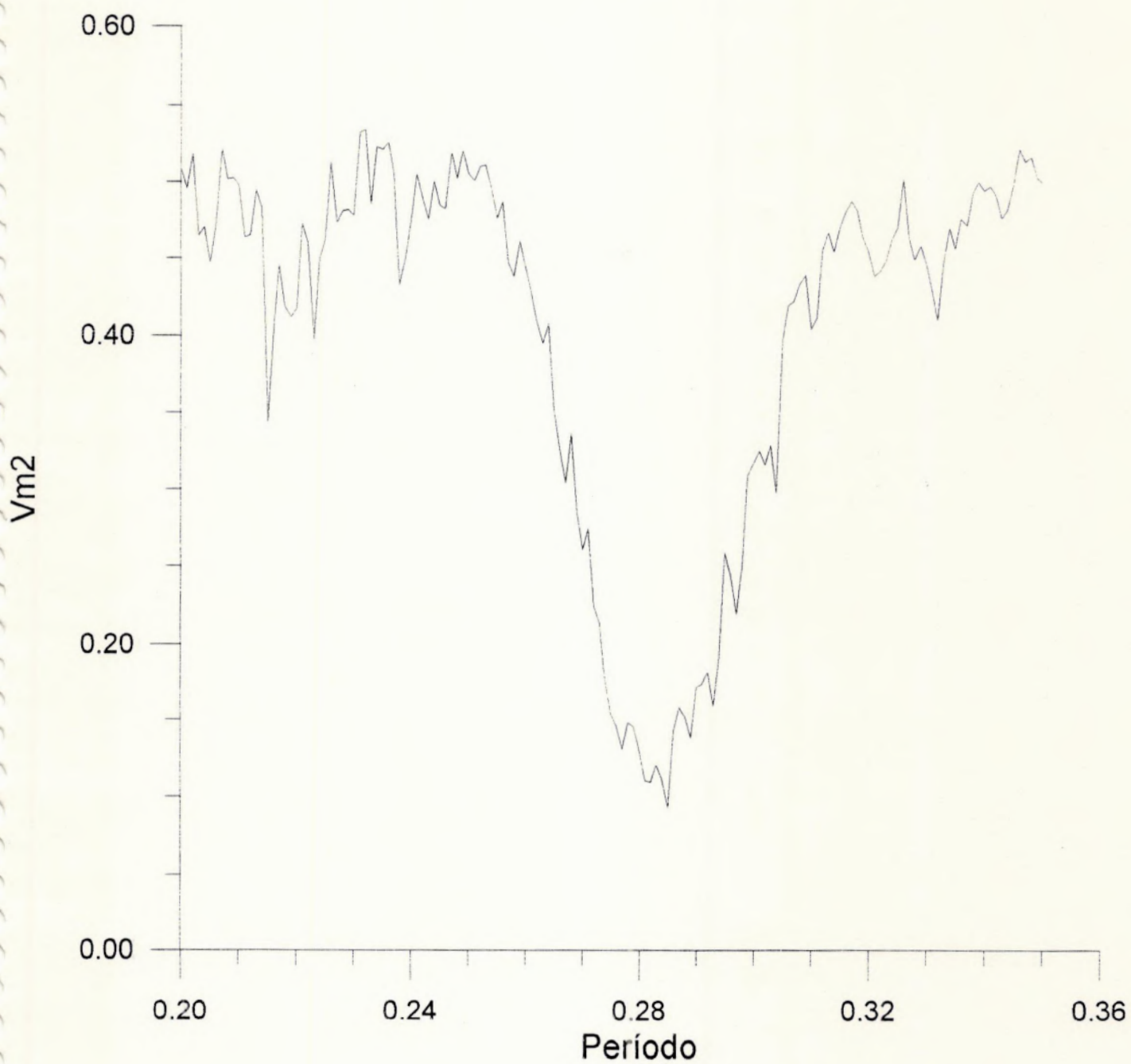


Fig 8- Periodograma Jurkevich para precisão 0.001

(Série VW CEPHEI- 10 grupos)

Vmz

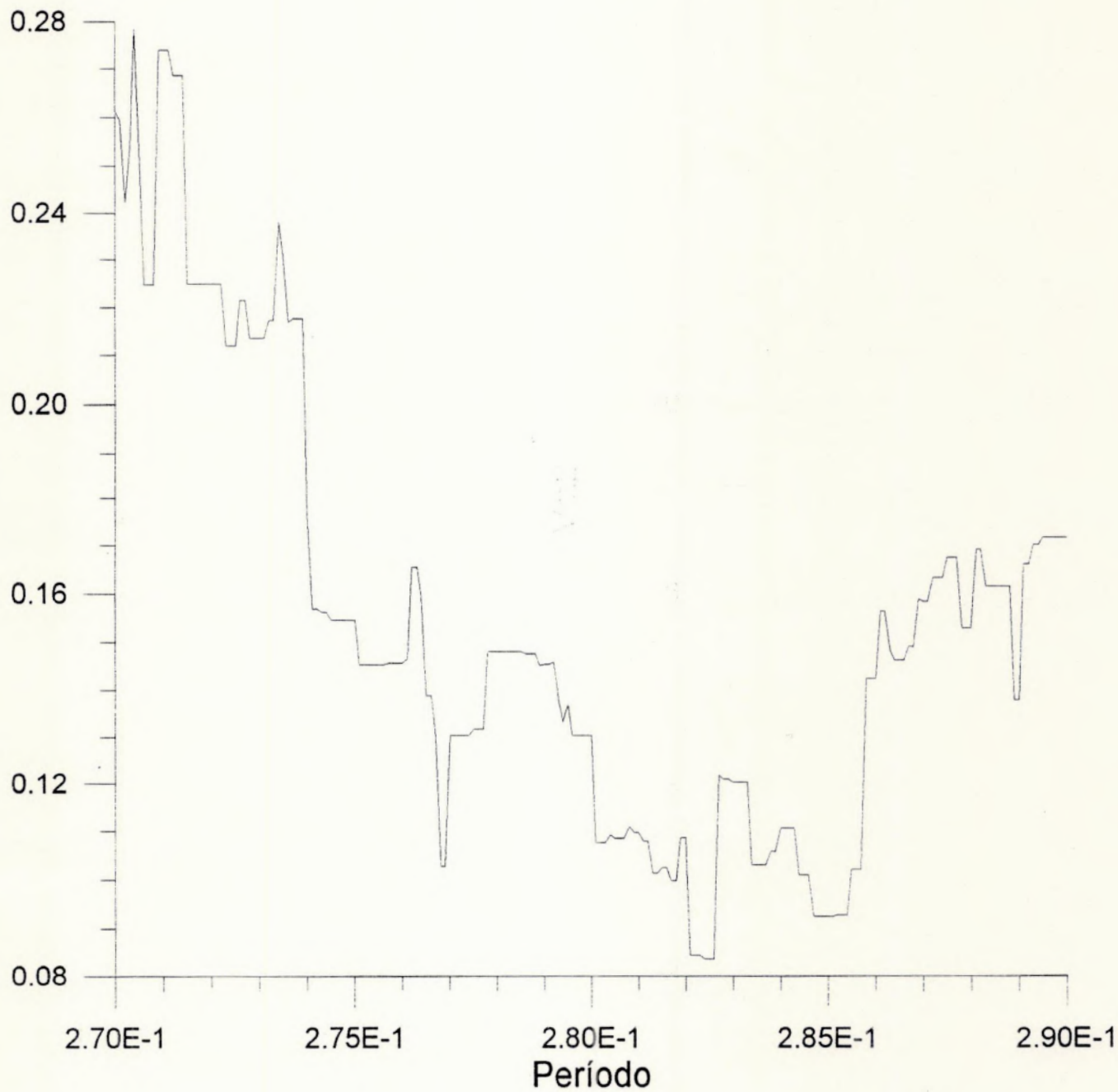


Fig 9- Periodograma Jurkevich para precisão 0.0001

(com intervalo de 0.27 a 0.29 para 10 grupos- série VW CEPHEI)

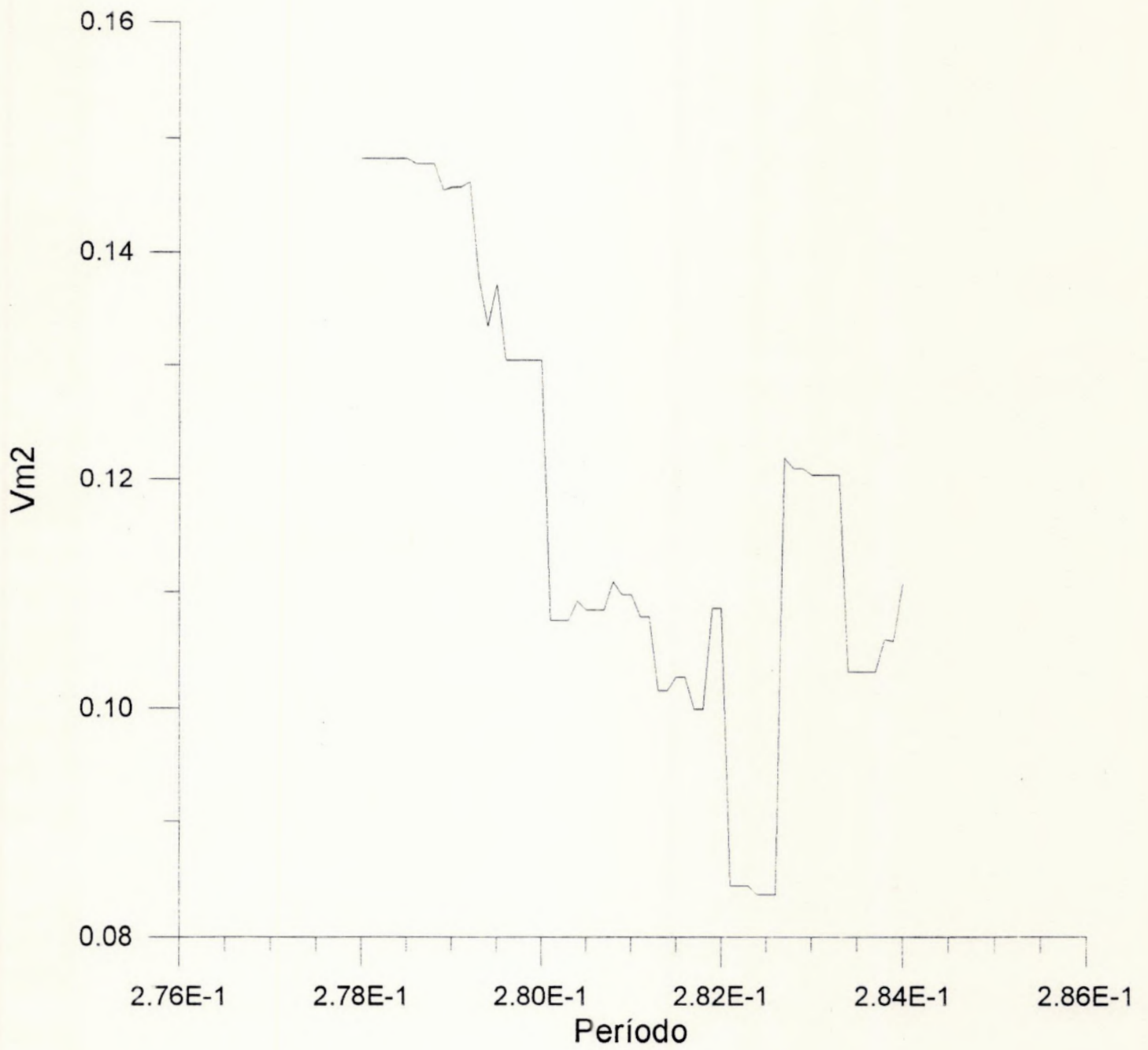


Fig. 10- Periodograma Jurkevich para precisão 0.0001

(intervalo 0.278 a 0.284)

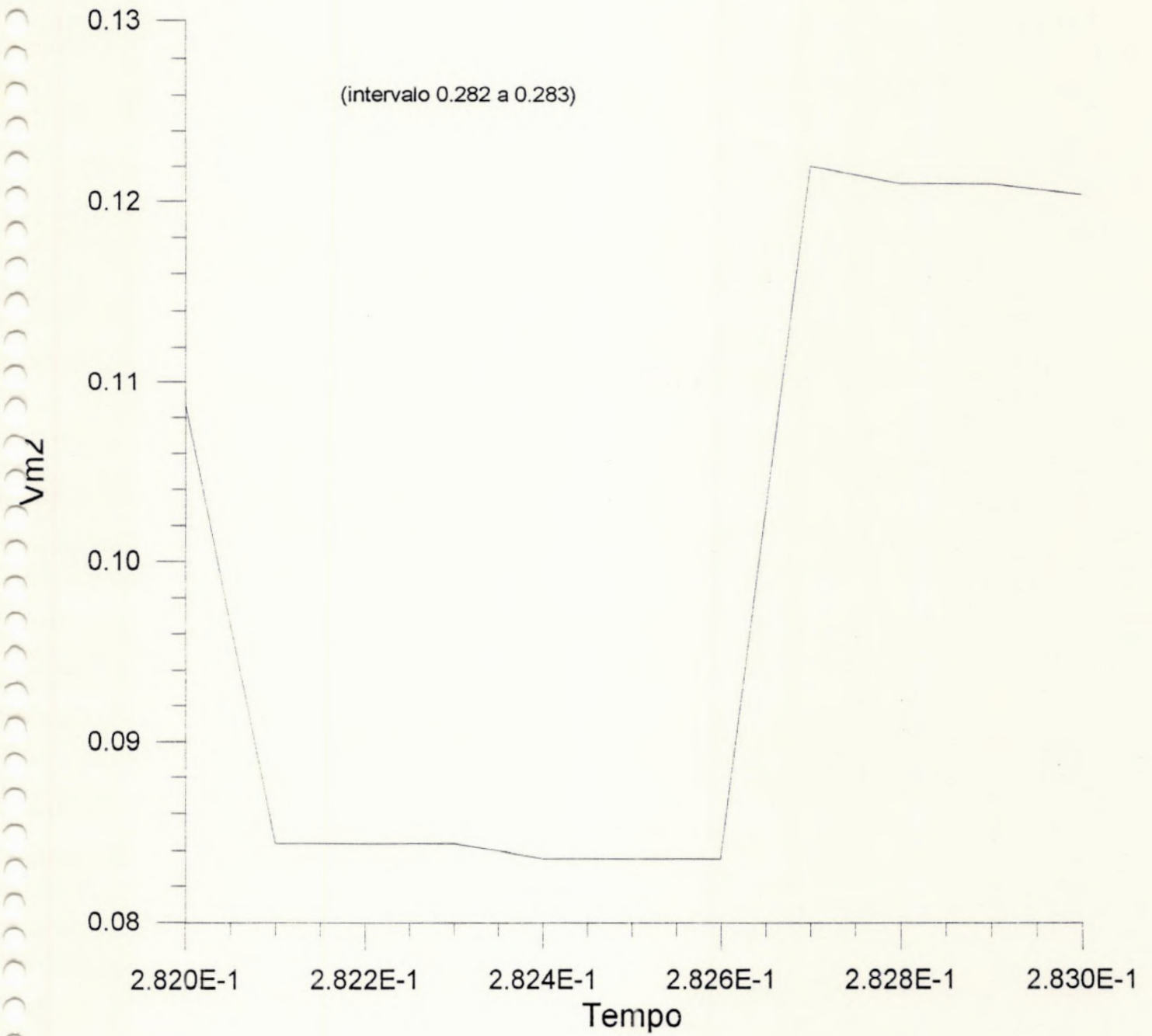


Fig 11- Periodograma Jurkevich para precisão 0.0001

Aproximação

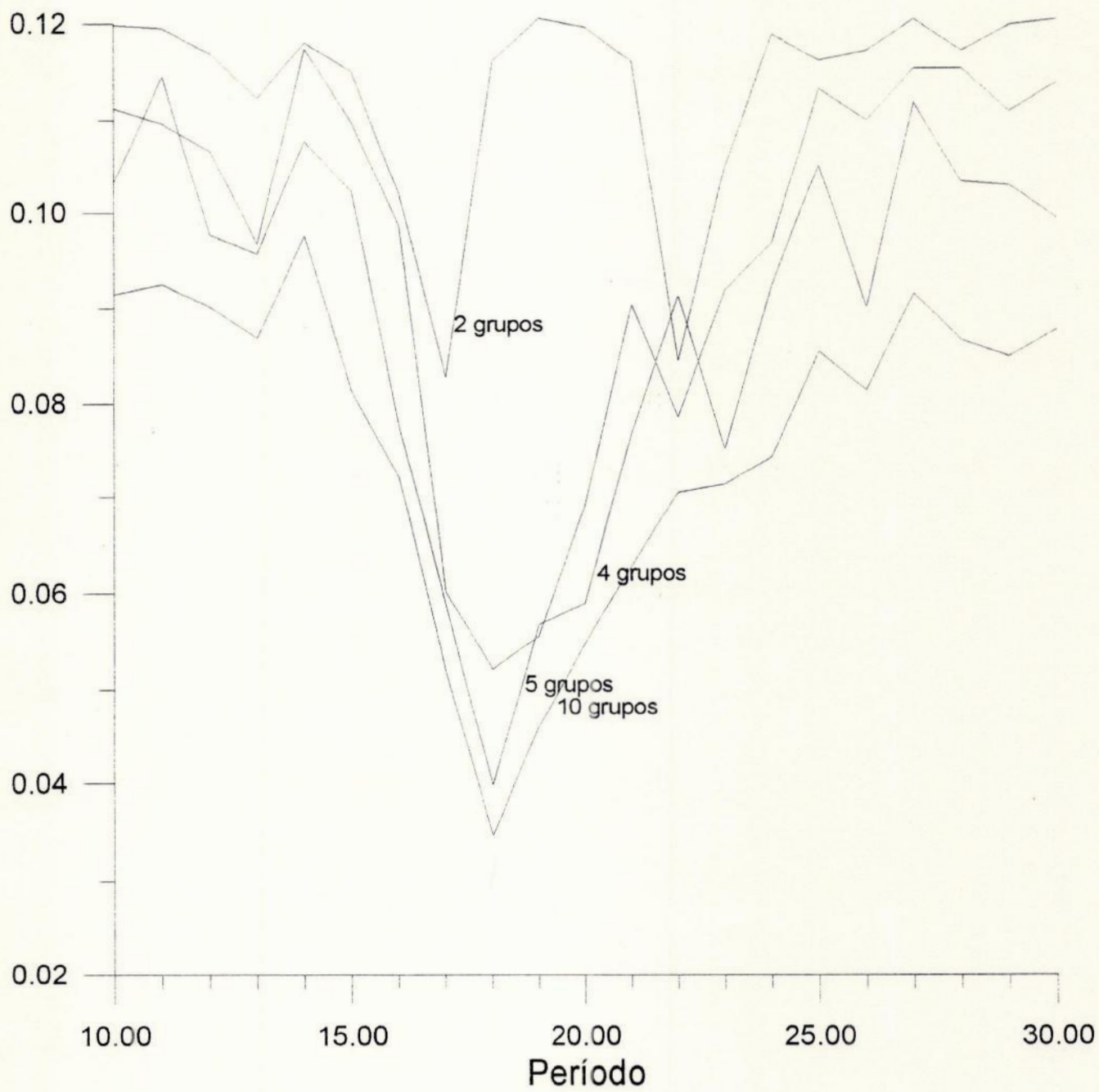


Fig. 12- Periodograma Jurkevich para 2, 4, 5 e 10 grupos
 (Para a série temporal sintética senoidal)

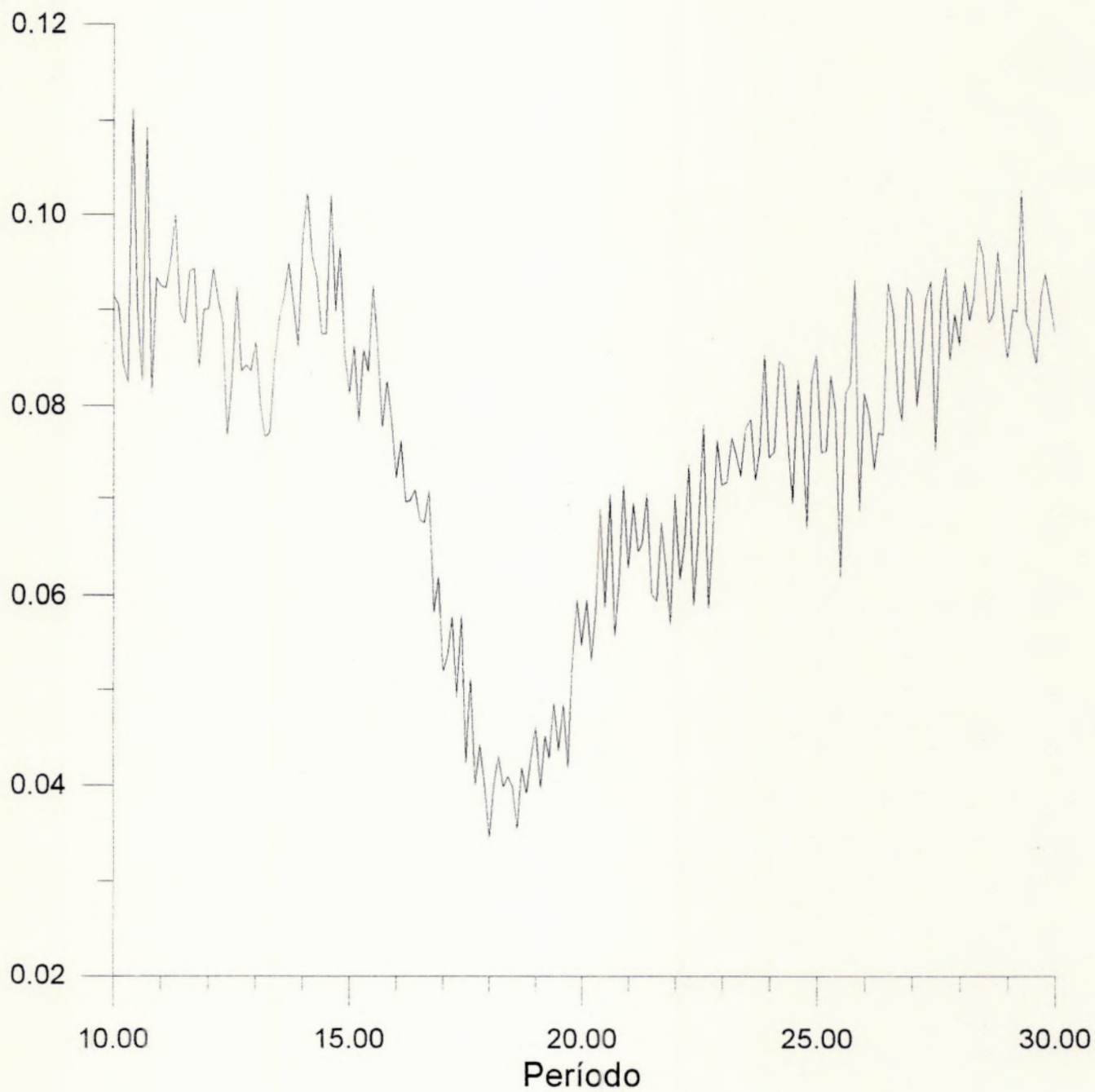


Fig. 13- Periodograma Jurkevich para precisão 0.1

(Intervalo de 10 a 30- série temporal sintética)

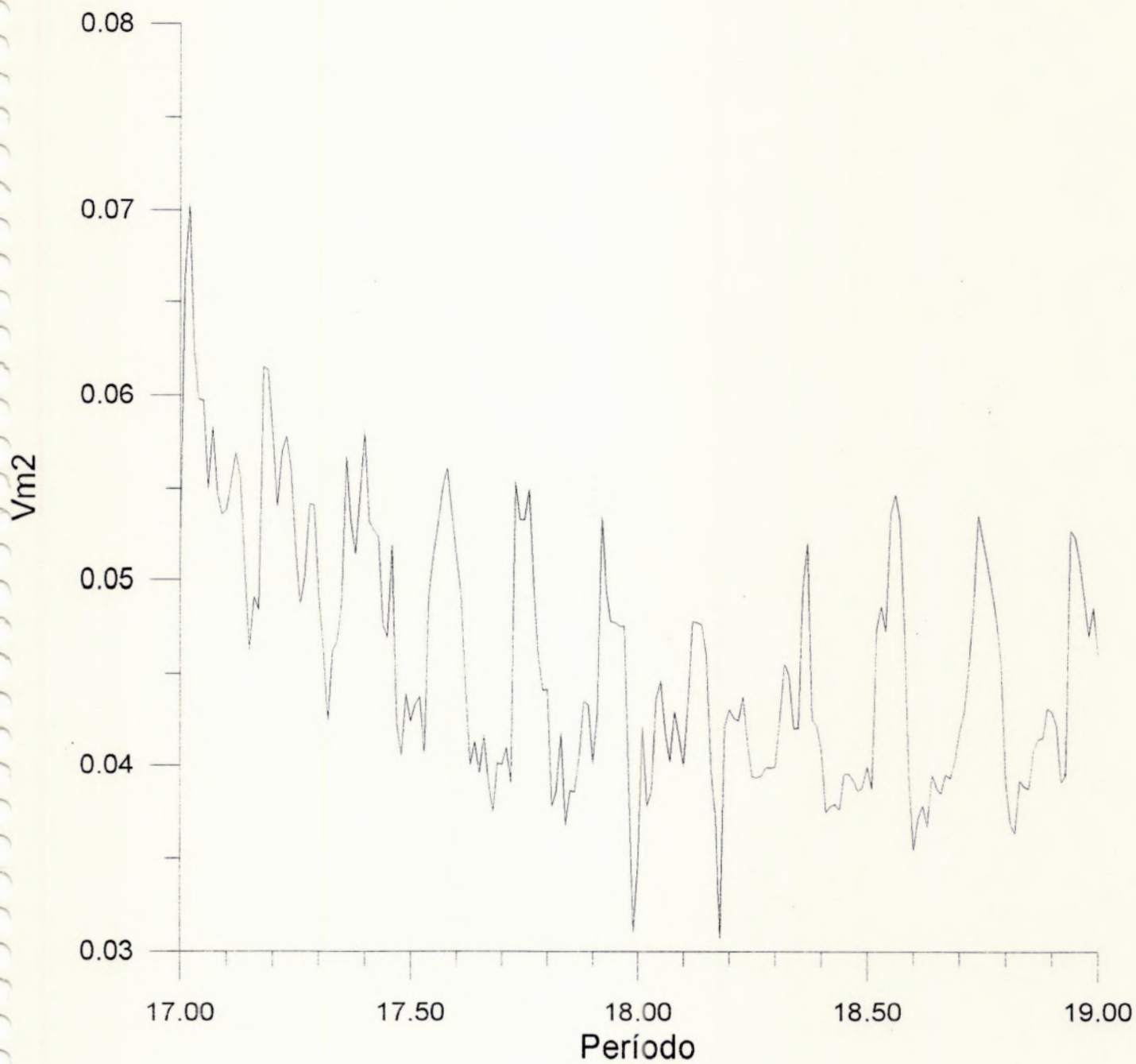


Fig. 14- Periodograma Jurkevich para precisão de 0.01

(intervalo de 17 a 19- série temporal sintética)

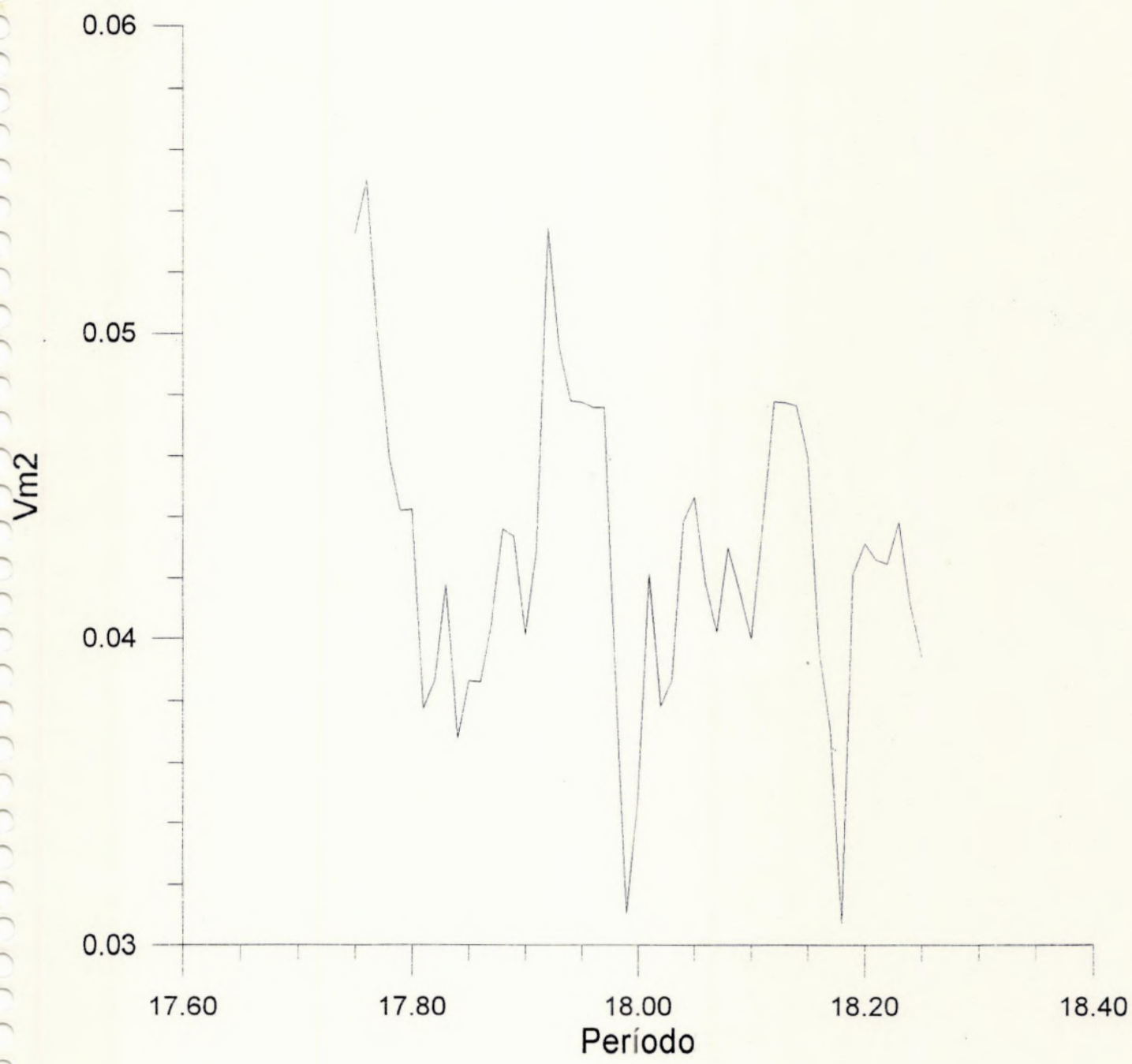


Fig. 15- Periodograma Jurkevich para precisão 0.01
 (intervalo 17.75 a 18.25- série temporal sintética)

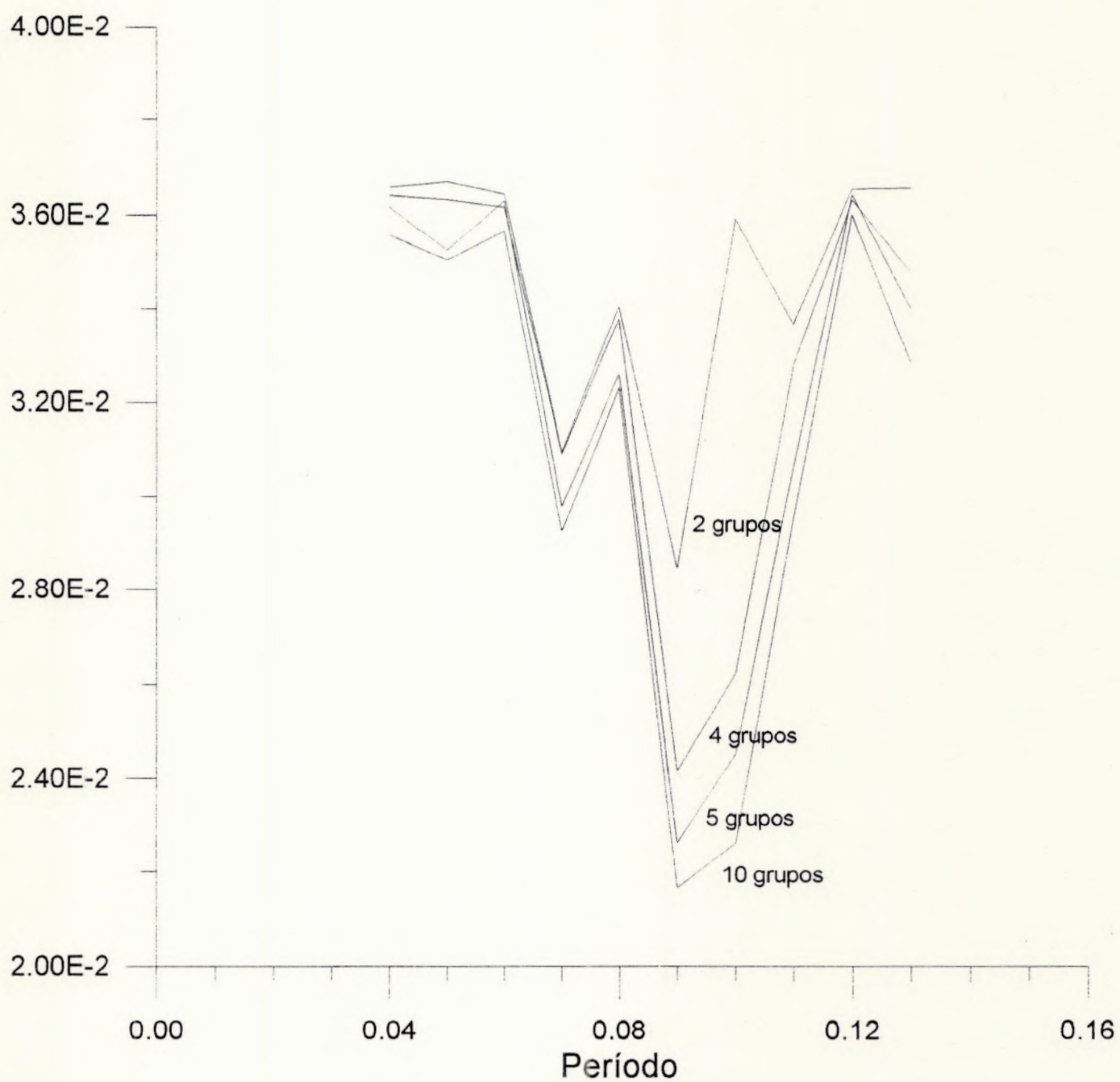


Fig. 16- Periodograma Jurkevich para 2, 4, 5, e 10 grupos

(Série temporal observada de HD93044)

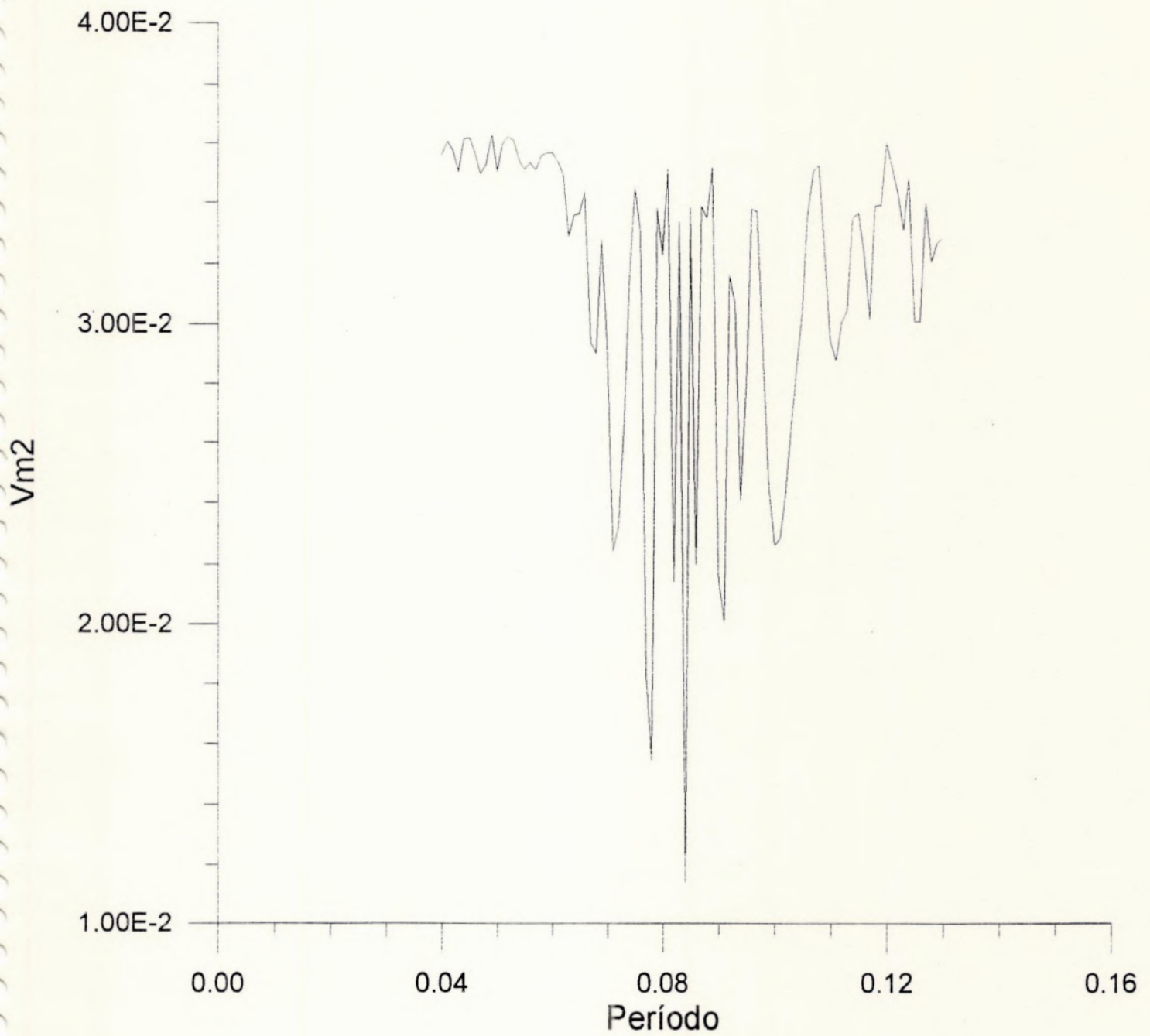


Fig. 17- Periodograma Jurkevich para precisão 0.001

(intervalo 0.04 a 0.13- HD93044)

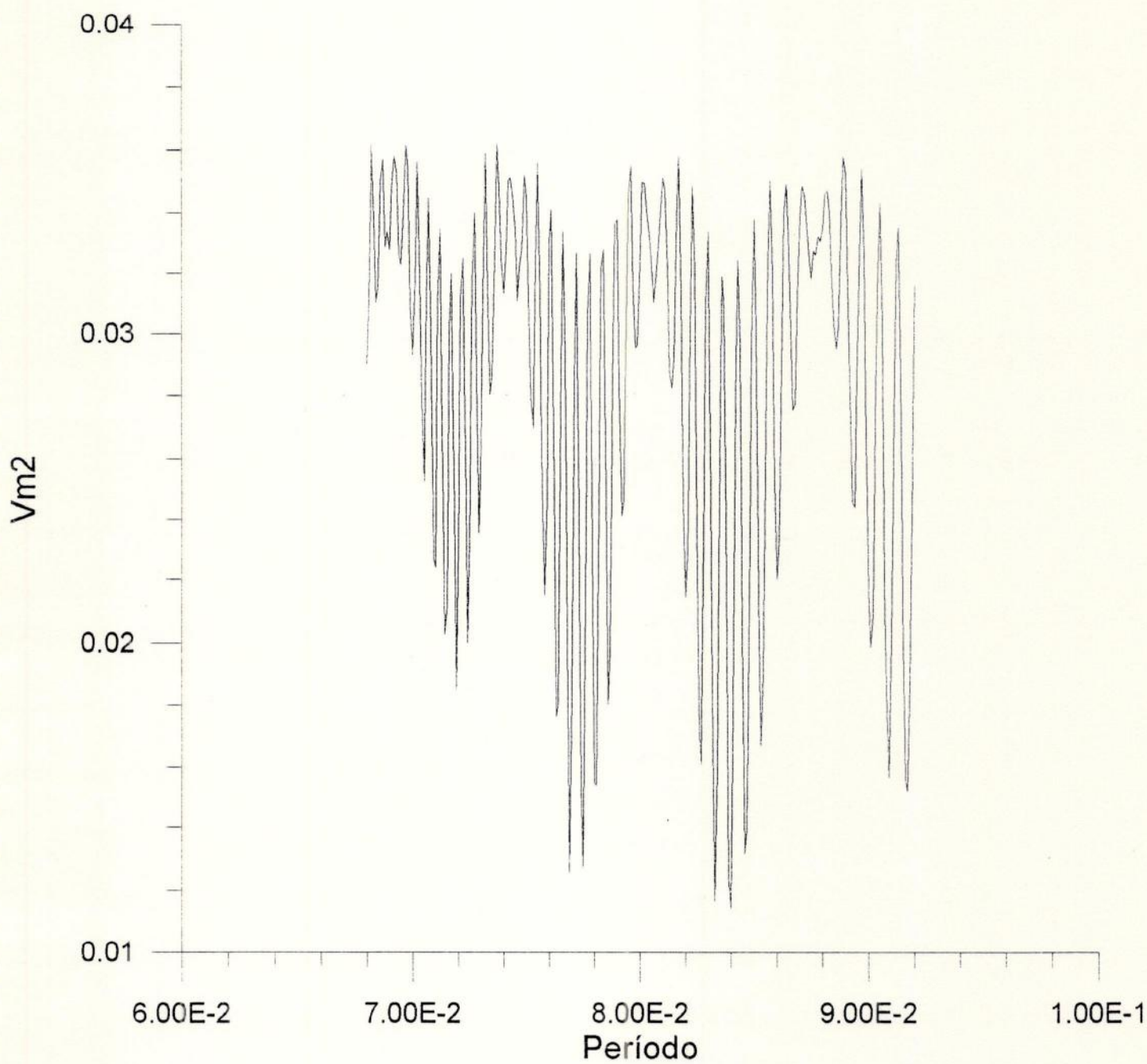


Fig. 18- Periodograma Jurkevich para precisão 0.0001

(intervalo 0.068 a 0.092- HD93044)

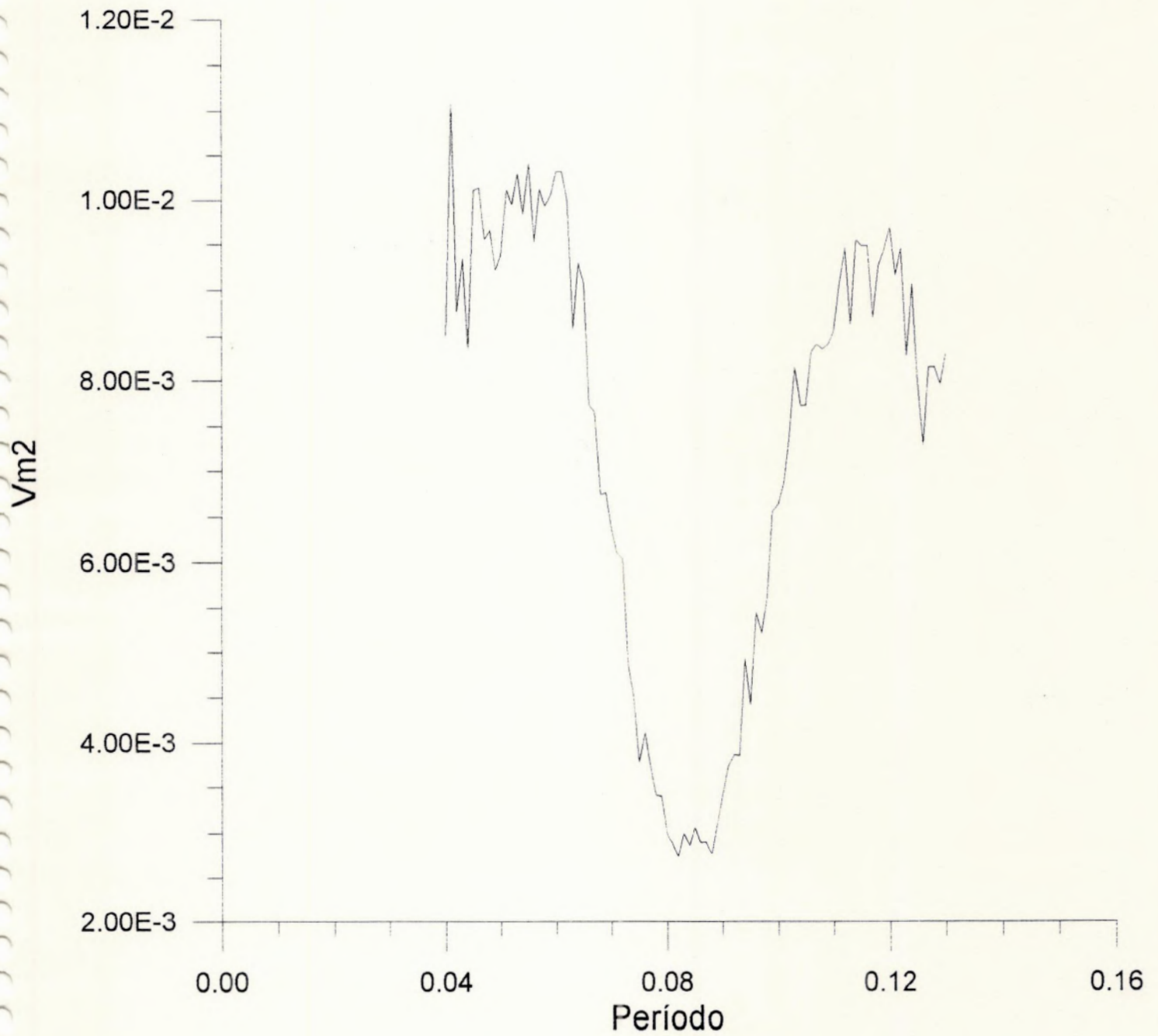


Fig. 19- Periodograma Jurkevich para Primeiro Conjunto de HD 93044

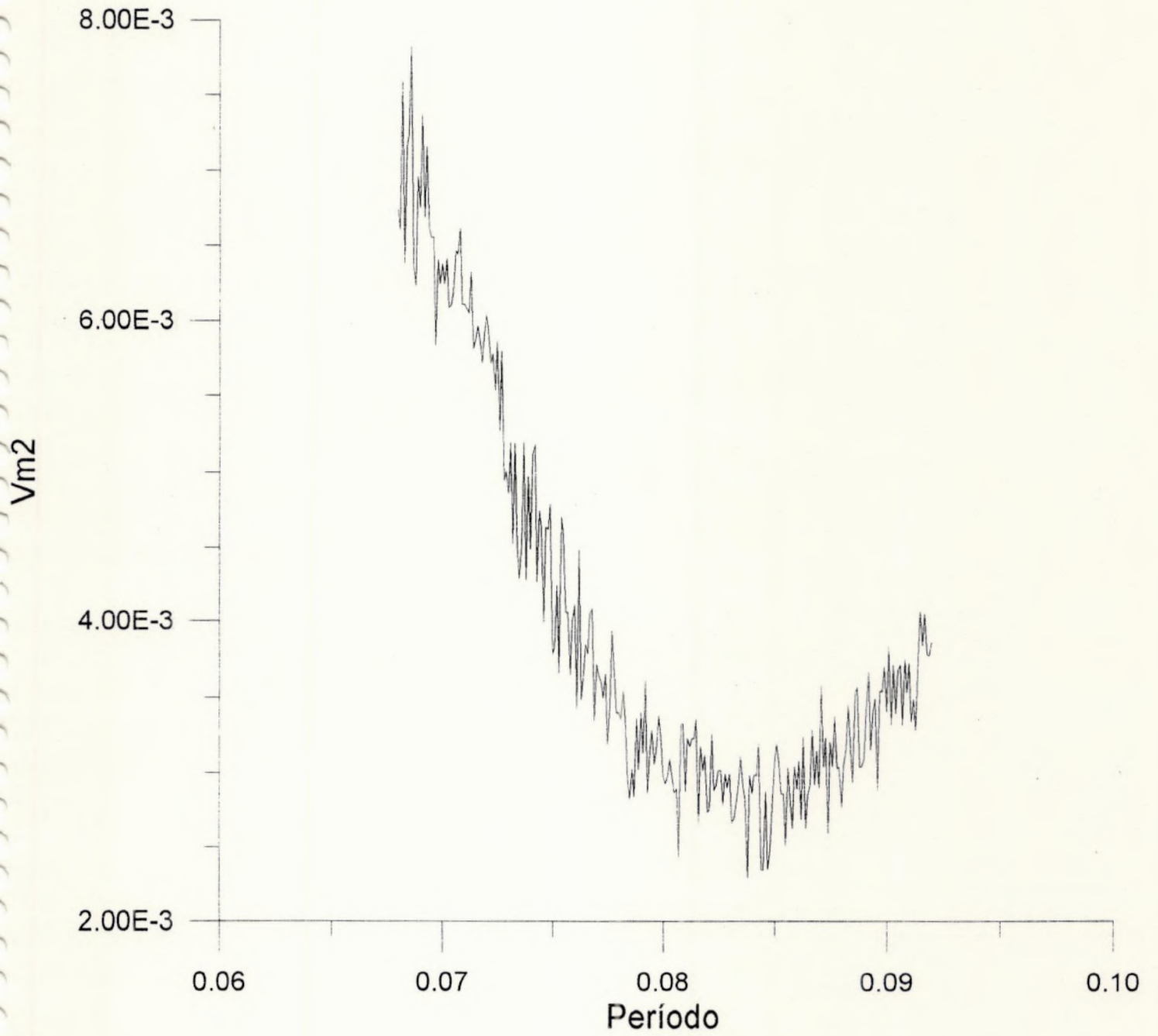


Fig. 20a- Periodograma Jurkevich para 0.068 a 0.092- Primeiro Conjunto de HD9304

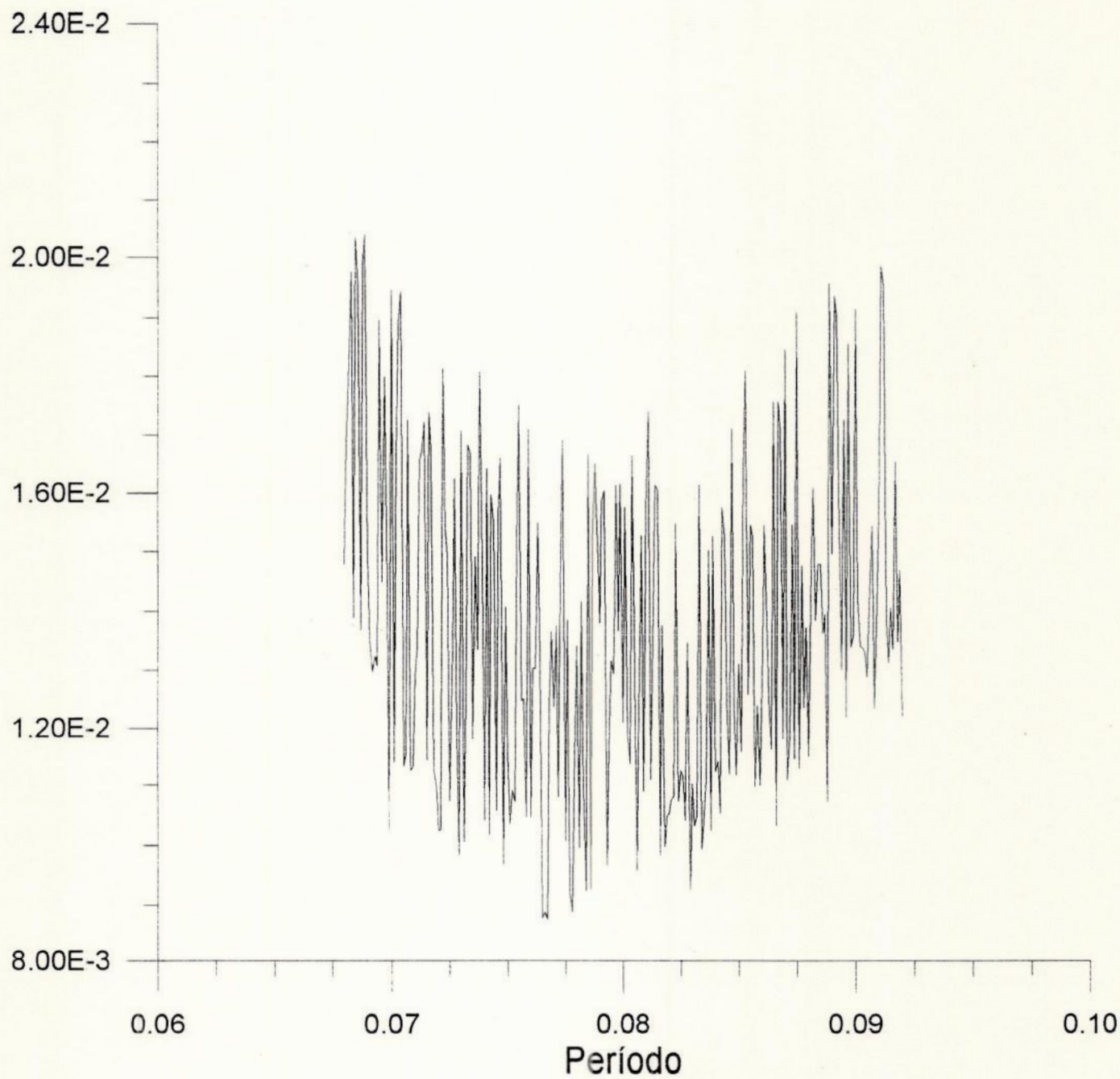


Fig. 20b- Periodograma Jurkevich para segundo conjunto de HD 93044

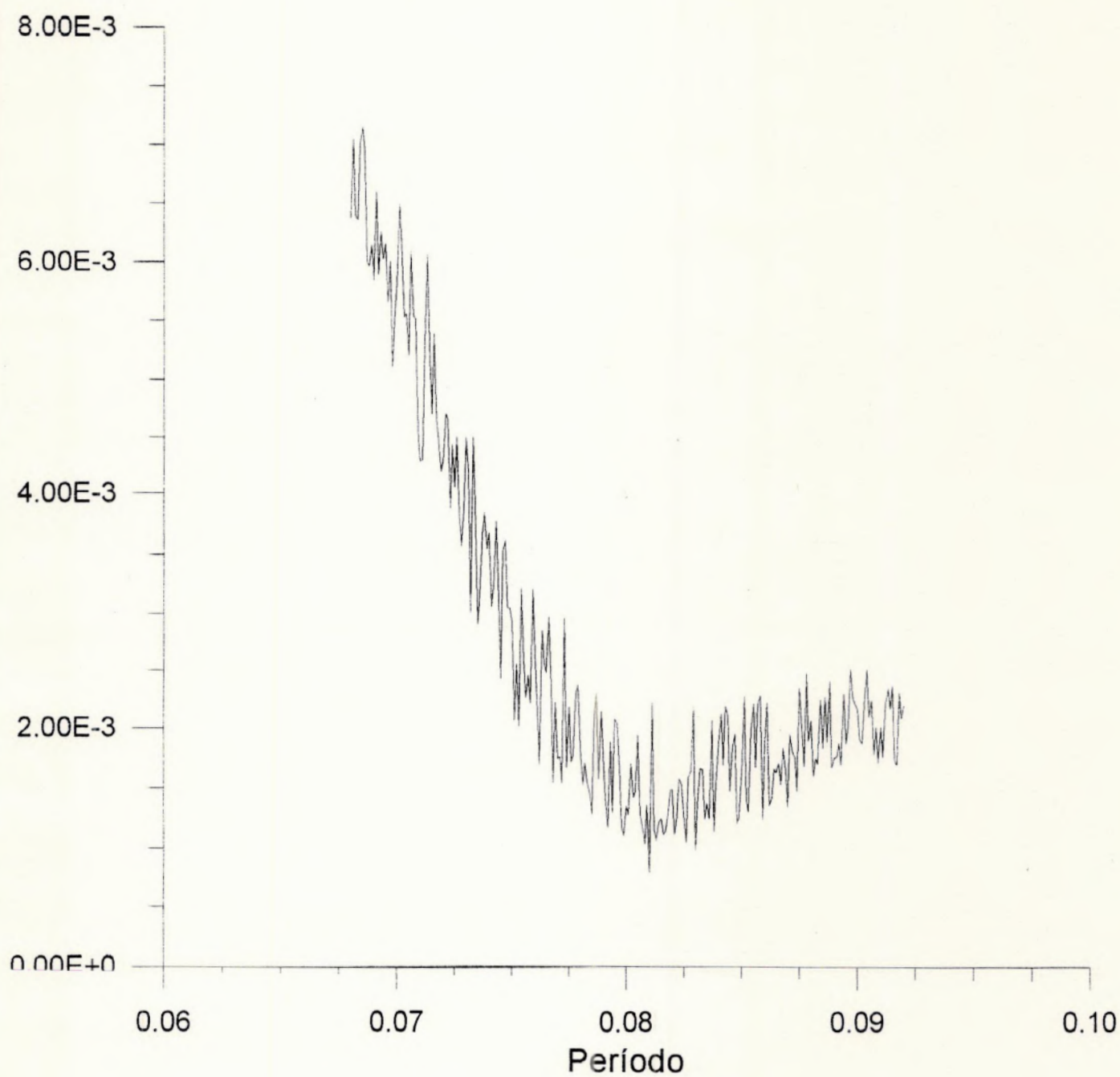


Fig 20c- Periodograma Jurkevich para terceiro conjunto de HD 93044

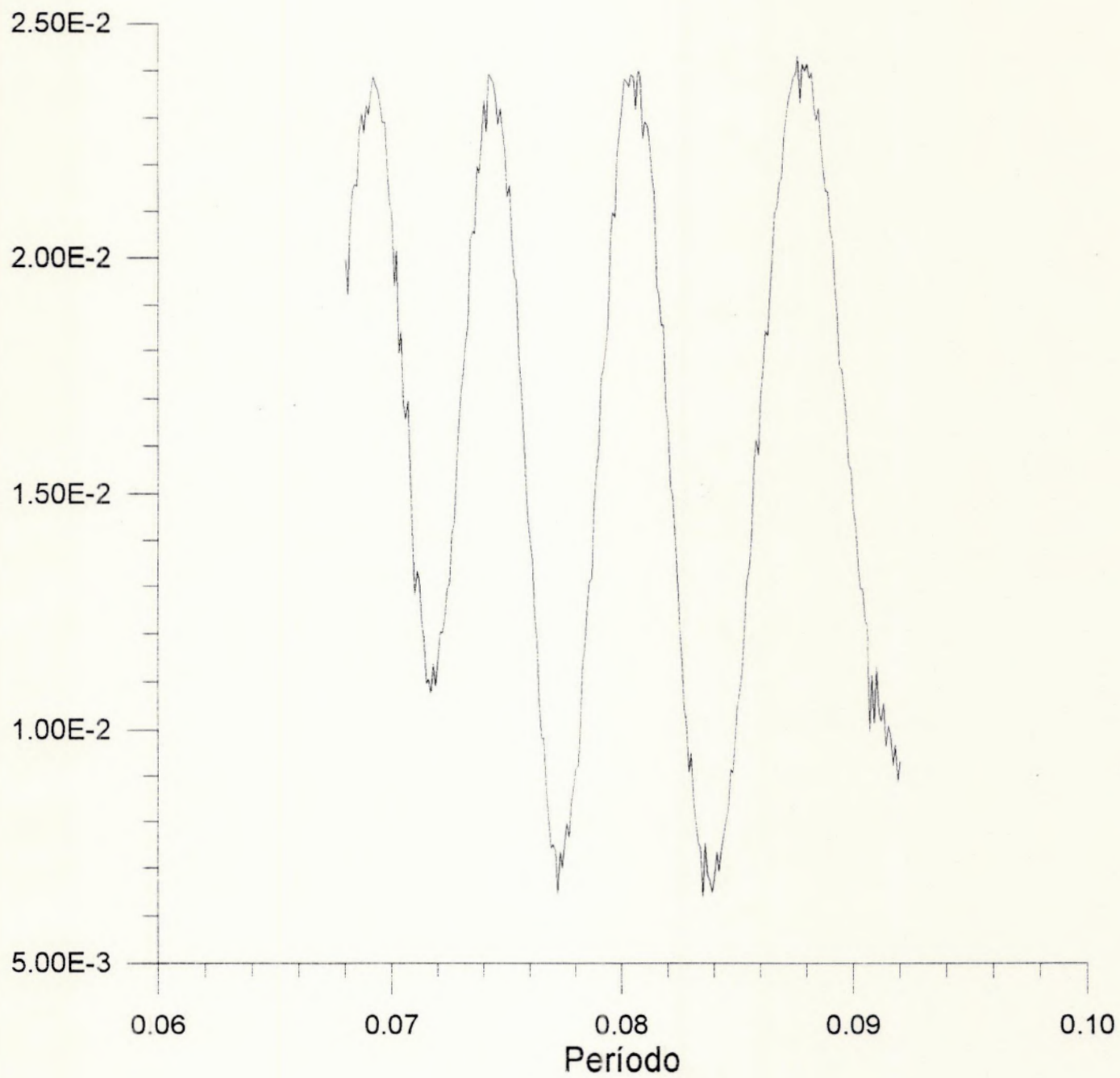


Fig 21- Periodogr Jurkevich para segundo e terceiro conjuntos de HD 93044 (0.068 a 0.092)

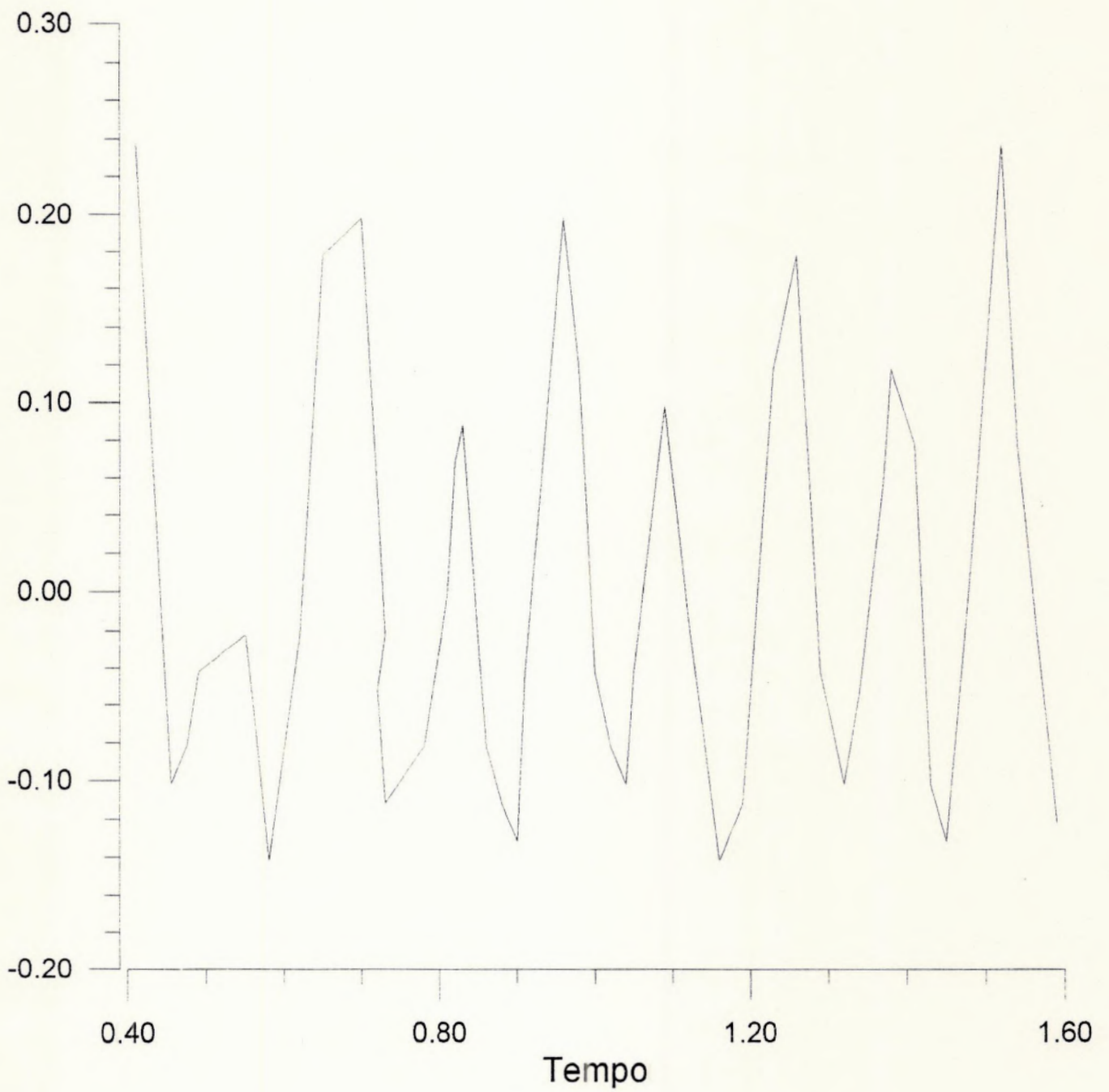


Fig. 22- Série Temporal de VW CEPHEI com média nula (subtraída)

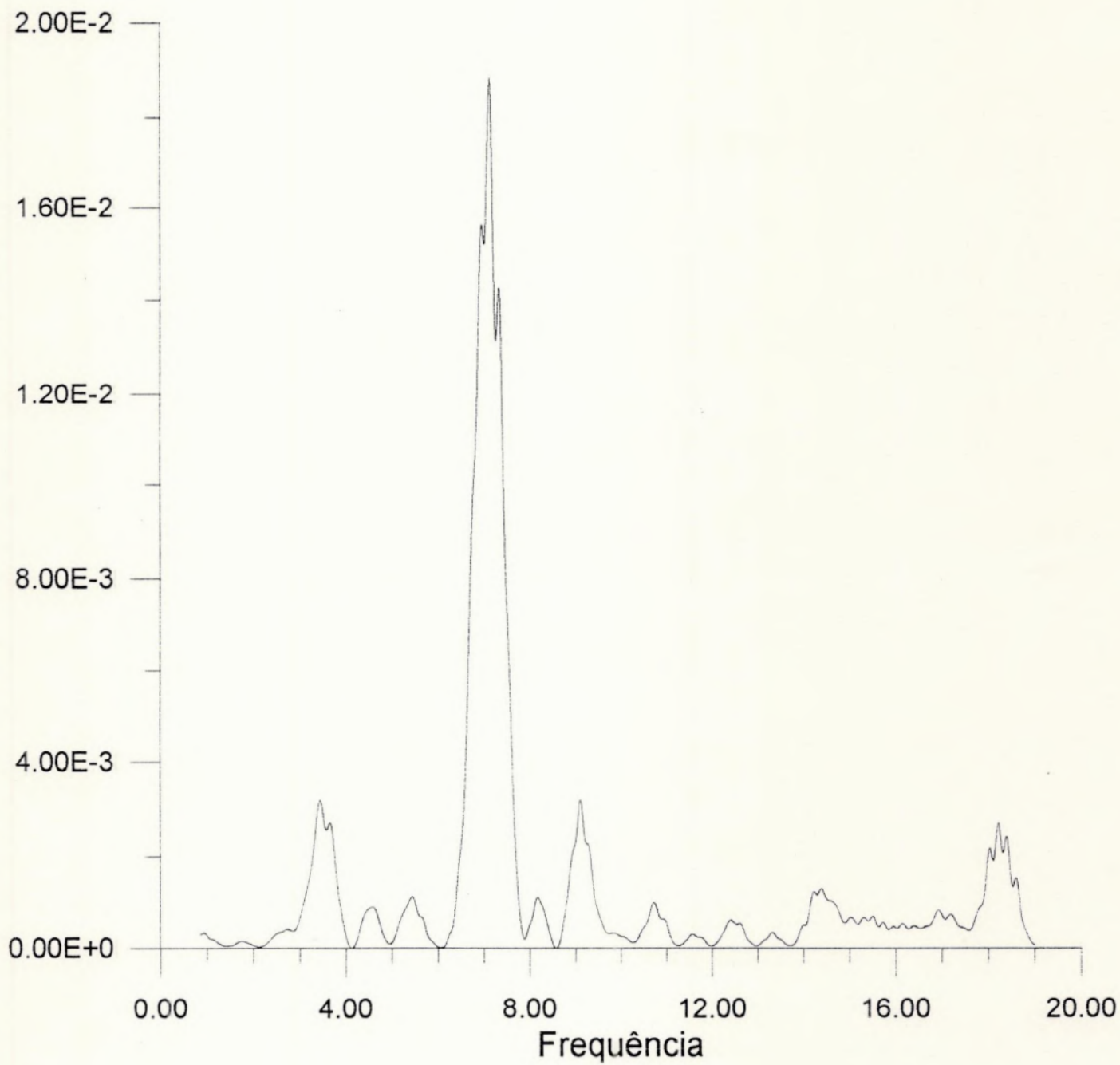


Fig. 23- Periodograma Ponman para VW CEPHEI

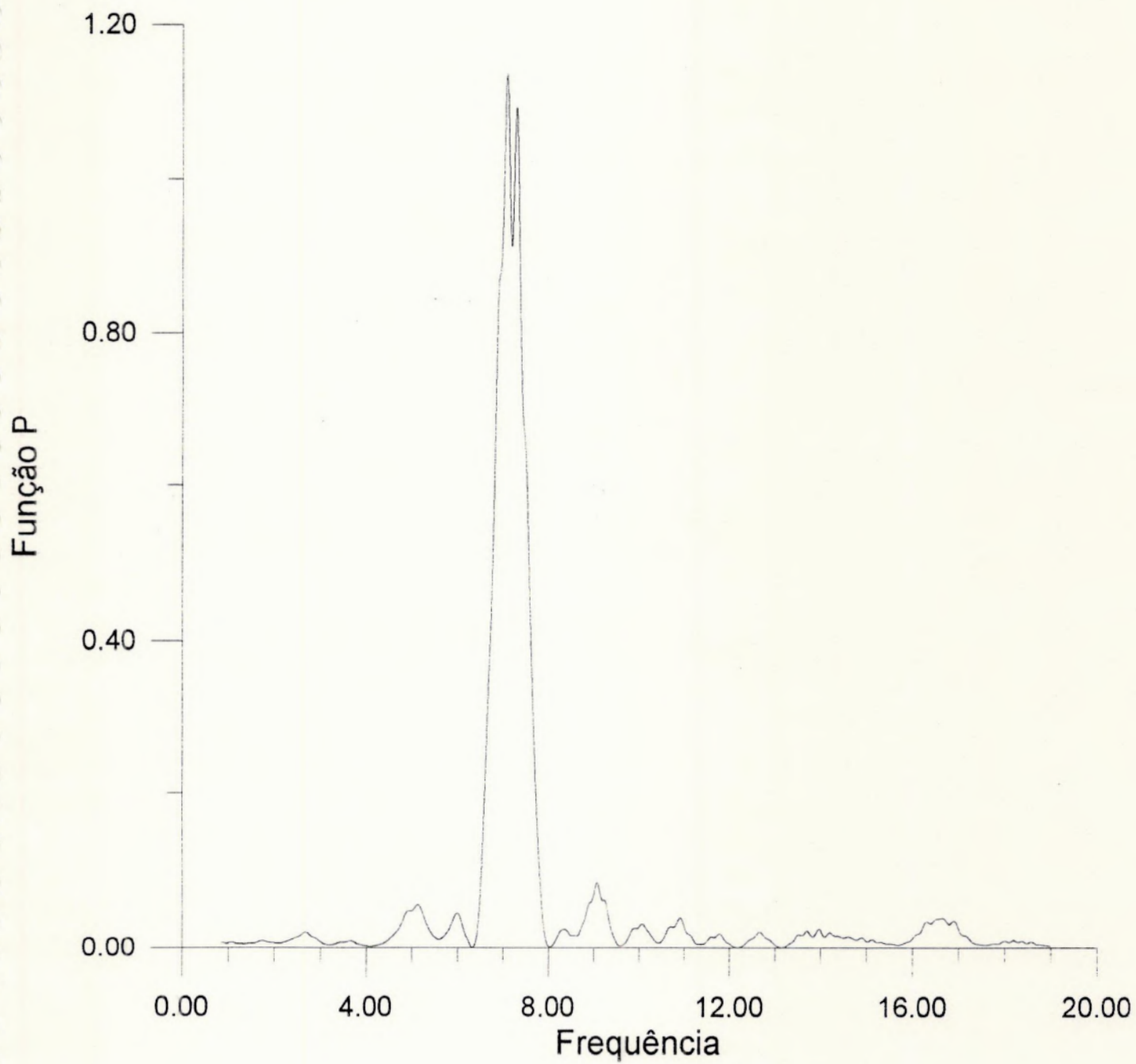


Fig 24- Periodograma Ponman para VW CEPHEI (Caso da Função Janela Senoidal)

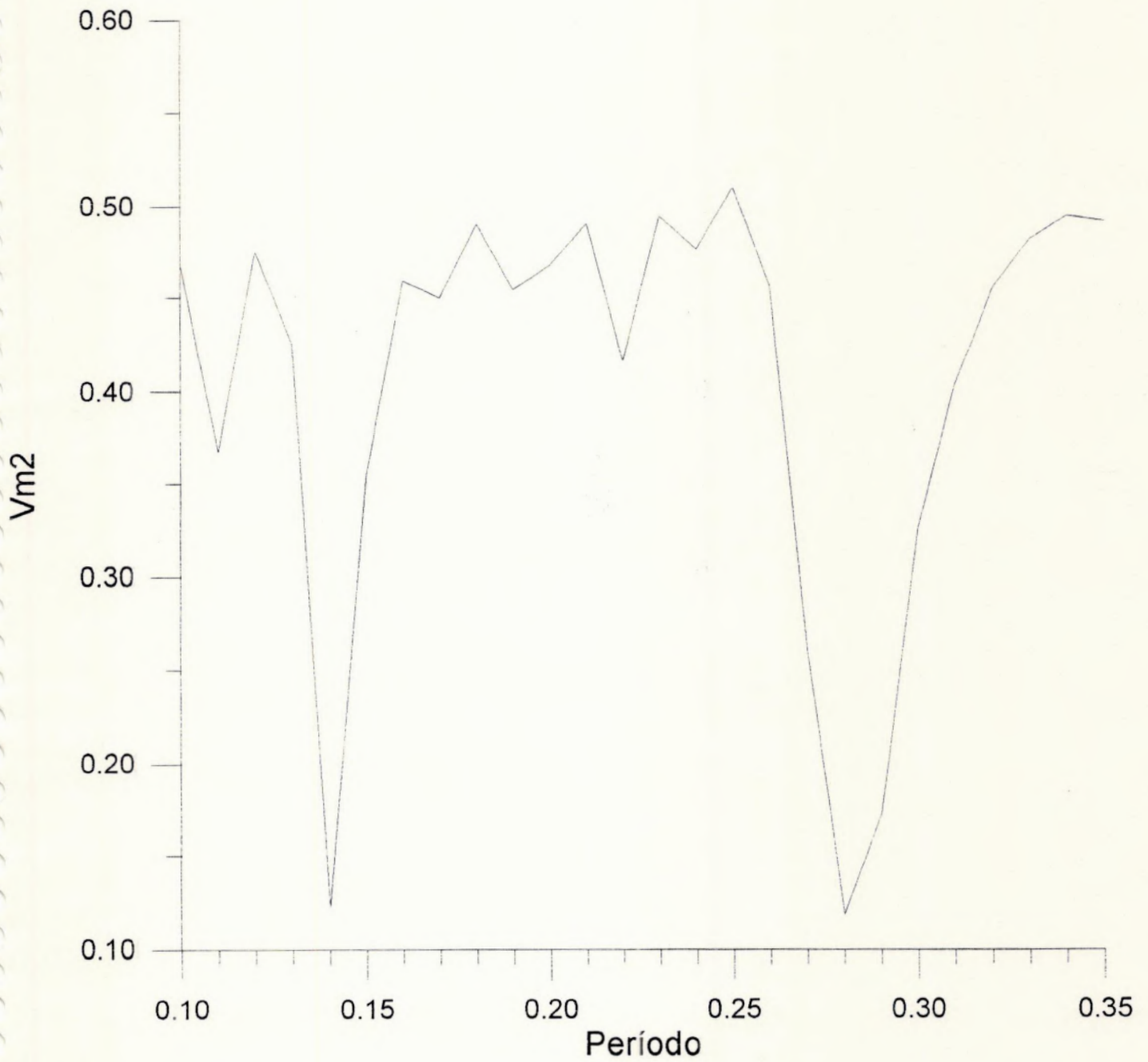


Fig 25- Periodograma Jurkevich revisto (Intervalo 0.10 a 0.35- VW CEPHEI- Precisão 0.01)

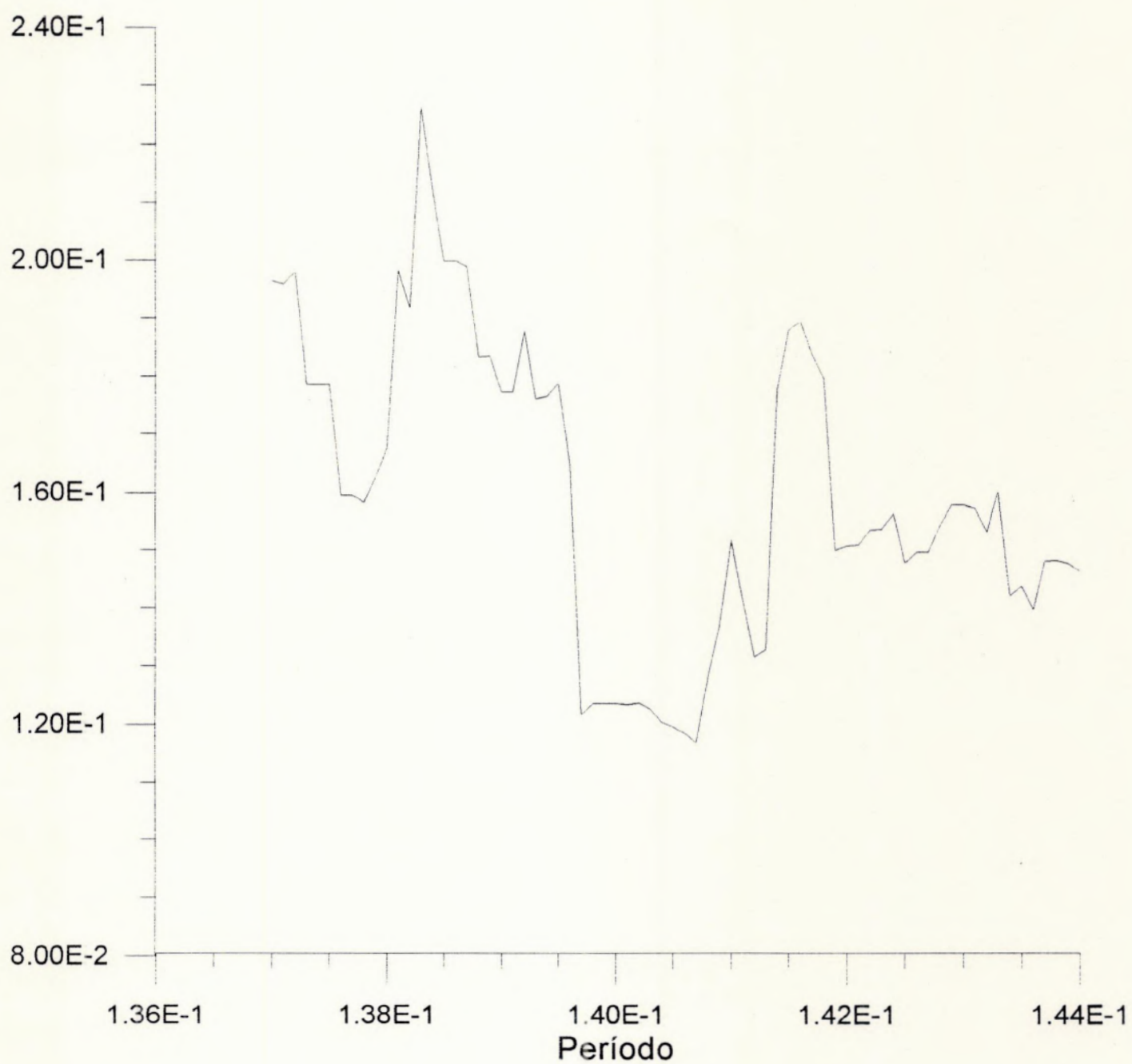


Fig. 25a- Periodograma Jurkevich revisto para VW CEPHEI (Intervalo 0.137 a 0.141)

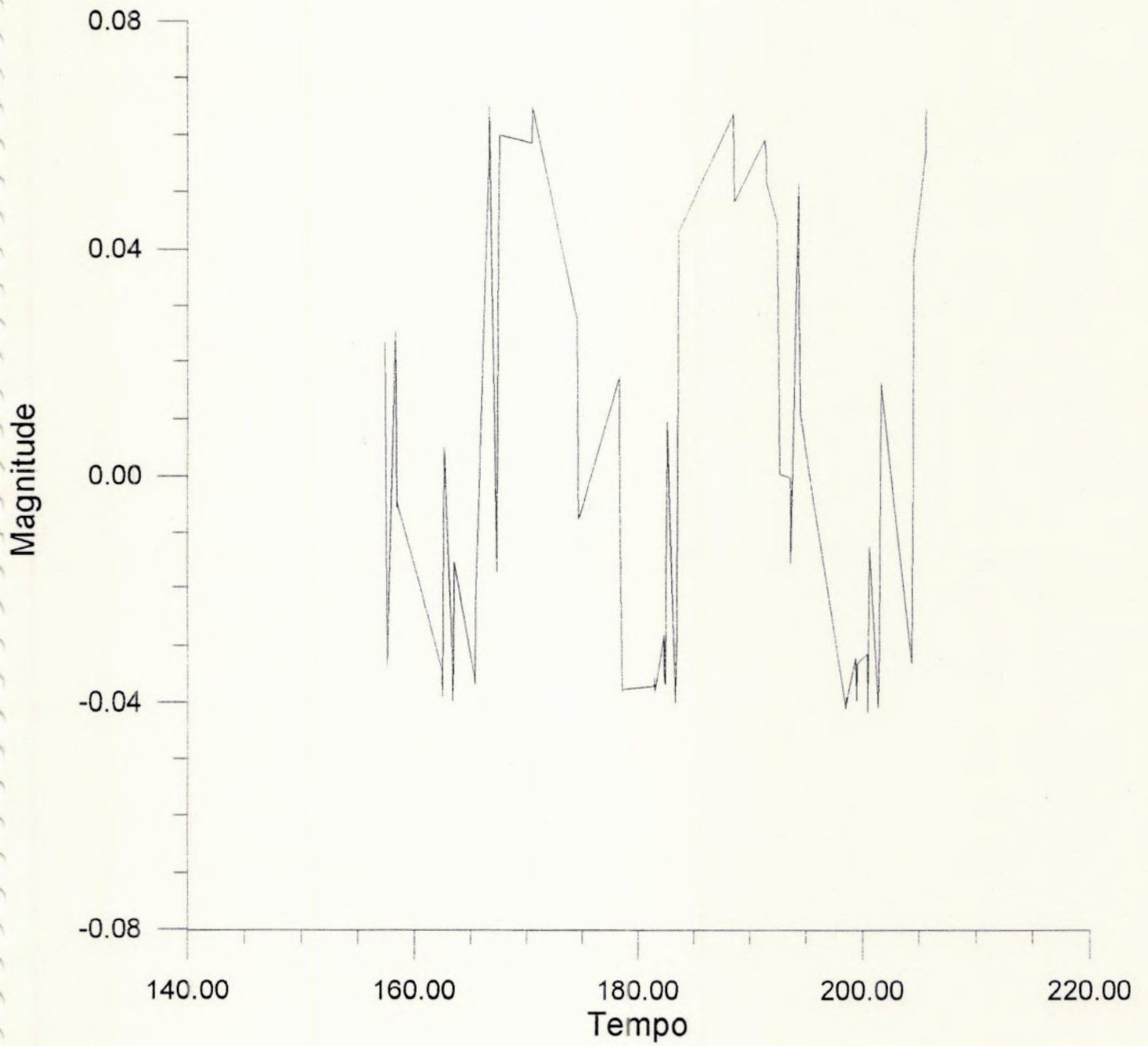


Fig. 26- Série Temporal Sintética Senoidal (Média Subtraída)

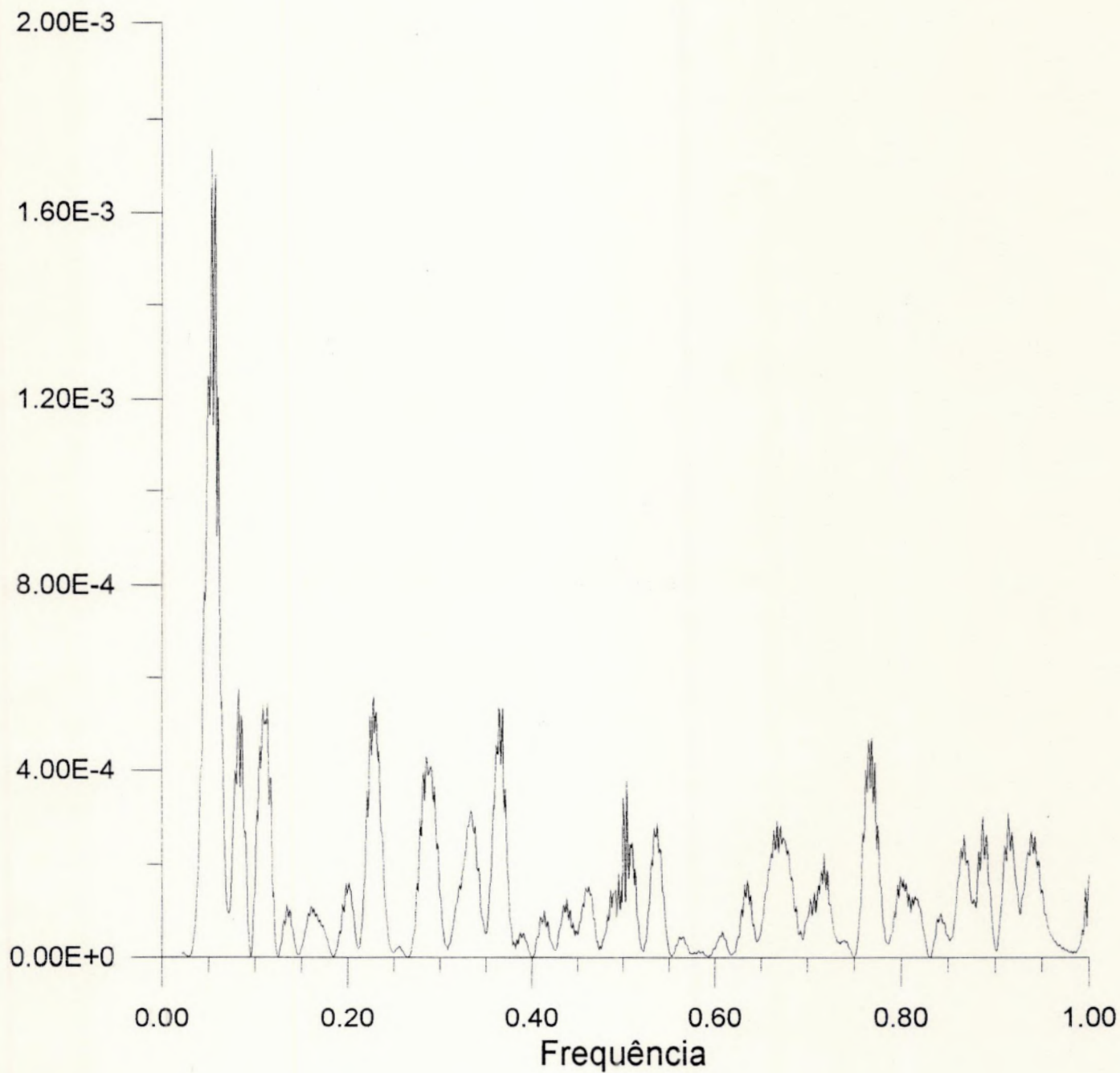


Fig 27- Periodograma Ponman para a série temporal sintética senoidal

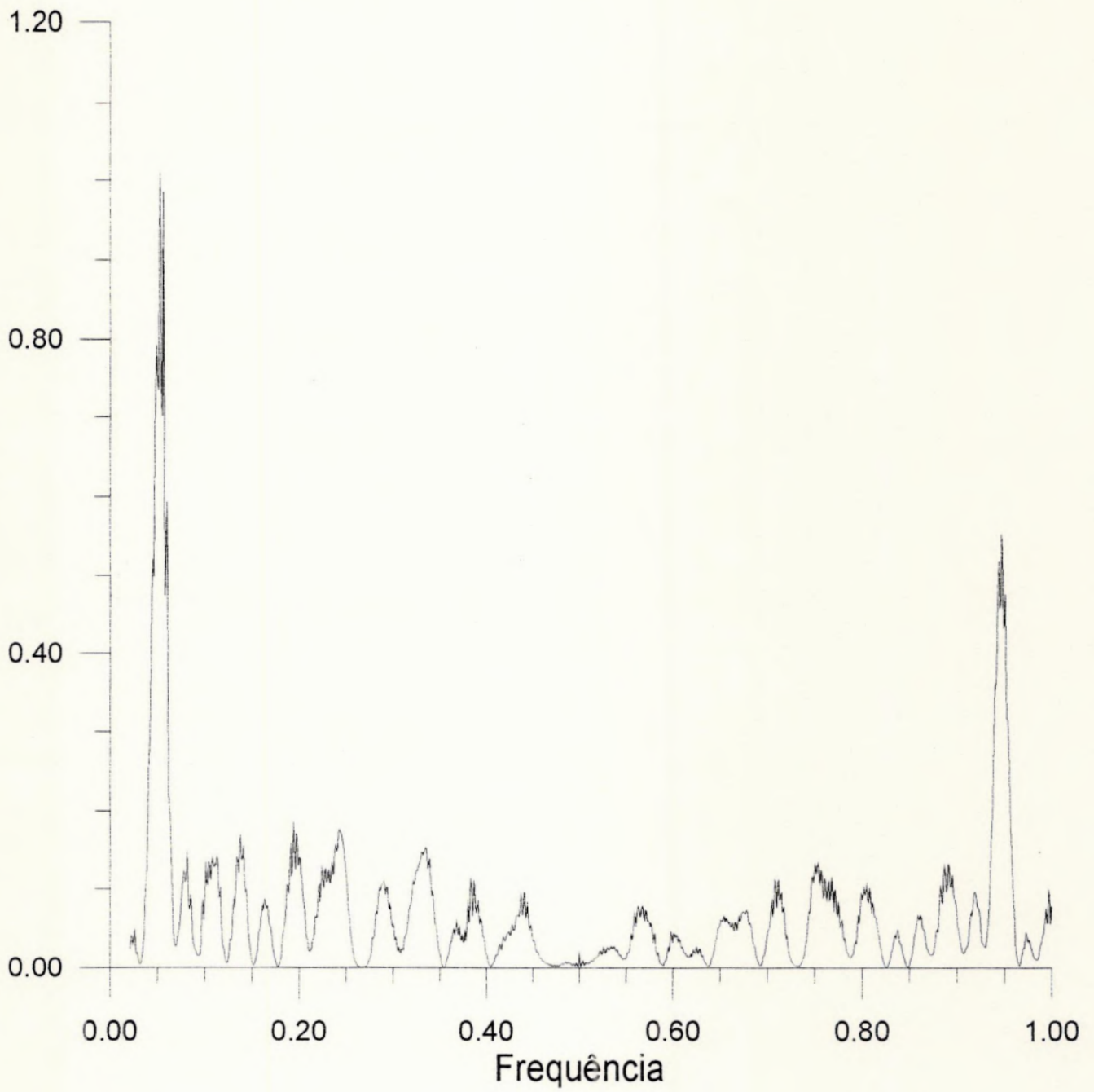


Fig 28- Periodograma Ponman para série temporal sintética senoidal (Caso da Função Janela Senoidal)

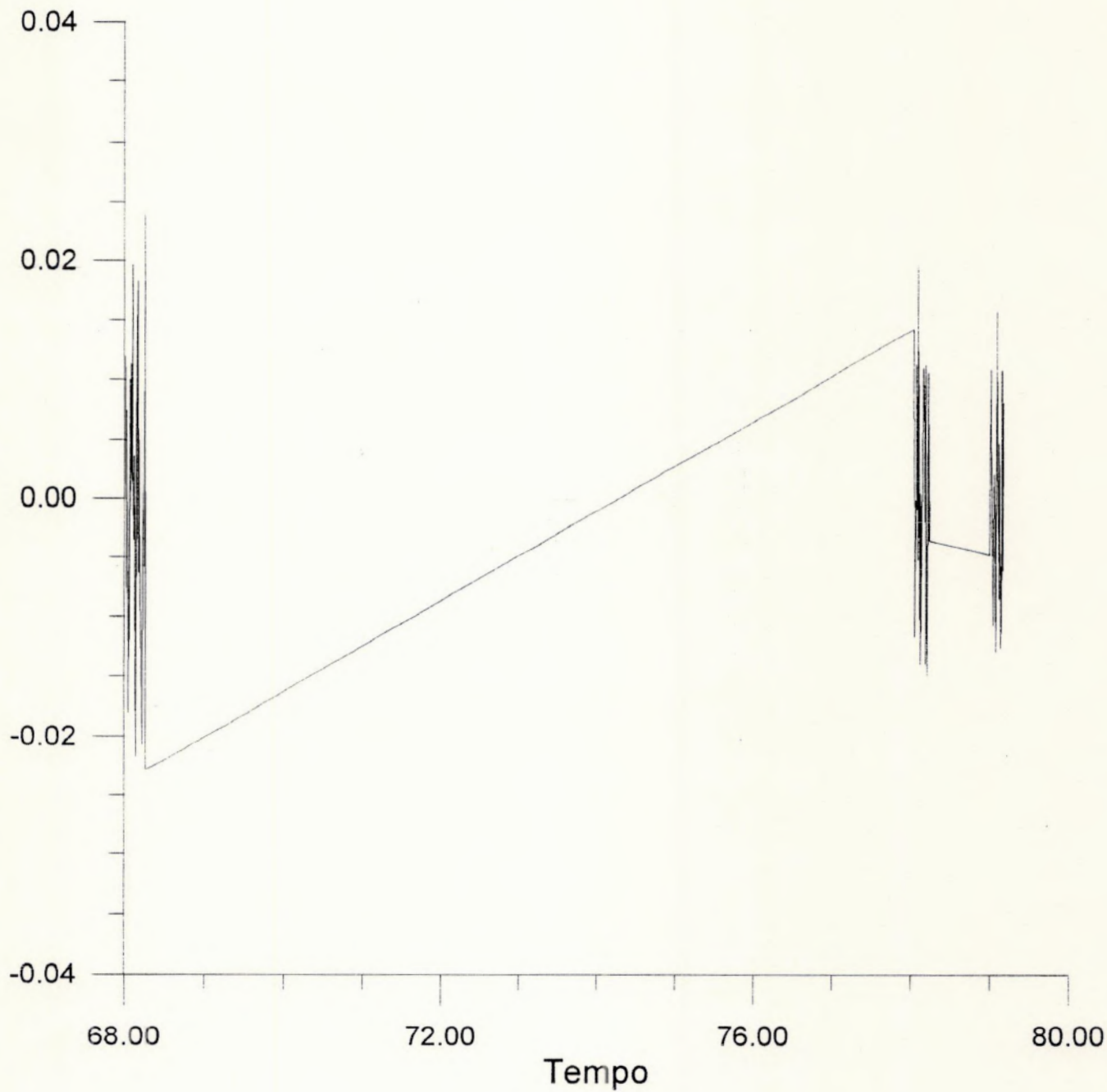


Fig. 29- Série Temporal de HD93044 com média nula e sem tendência lenta

(2448300 +)

Função F

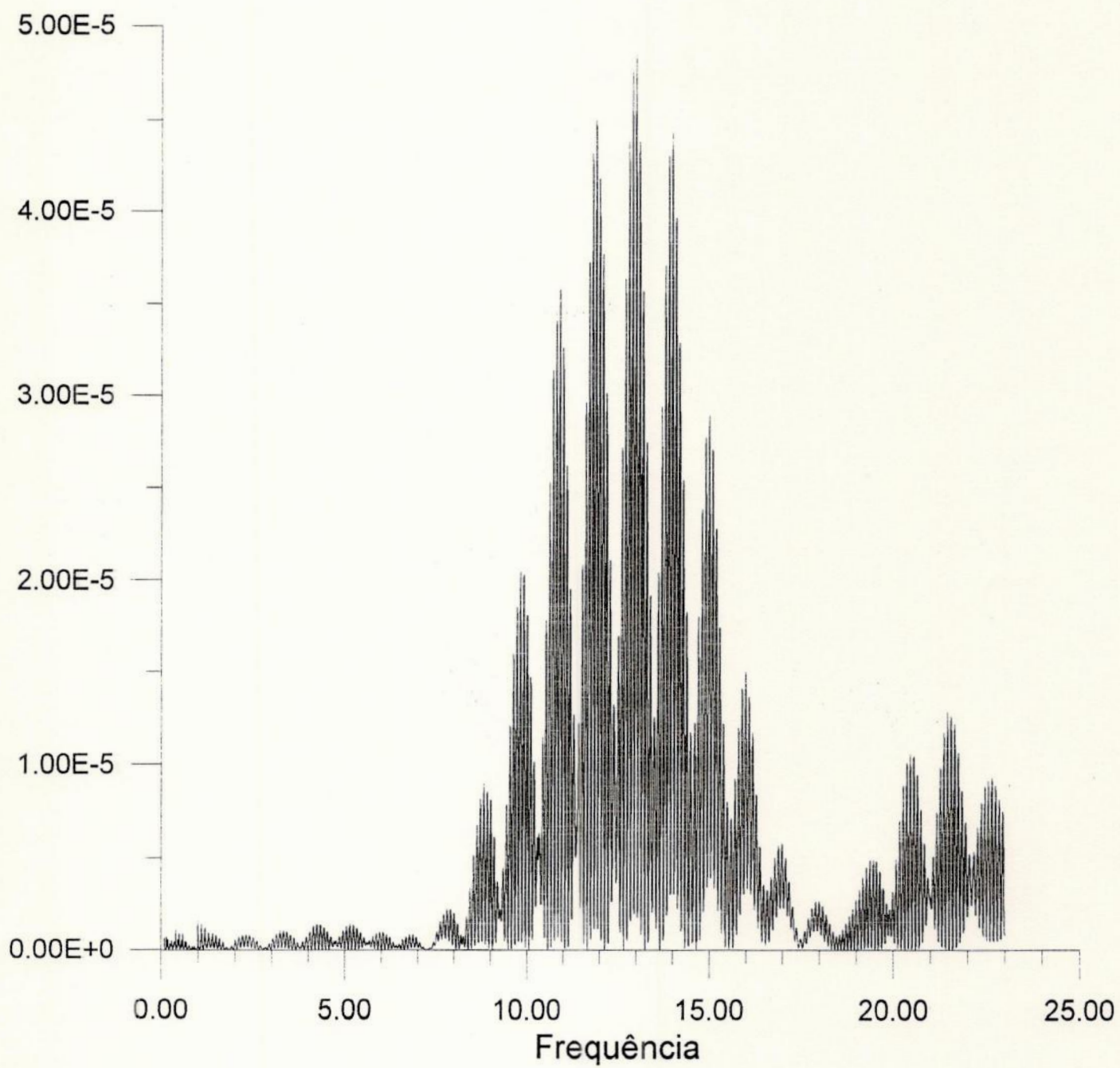


Fig 30- Periodograma Ponman para HD93044

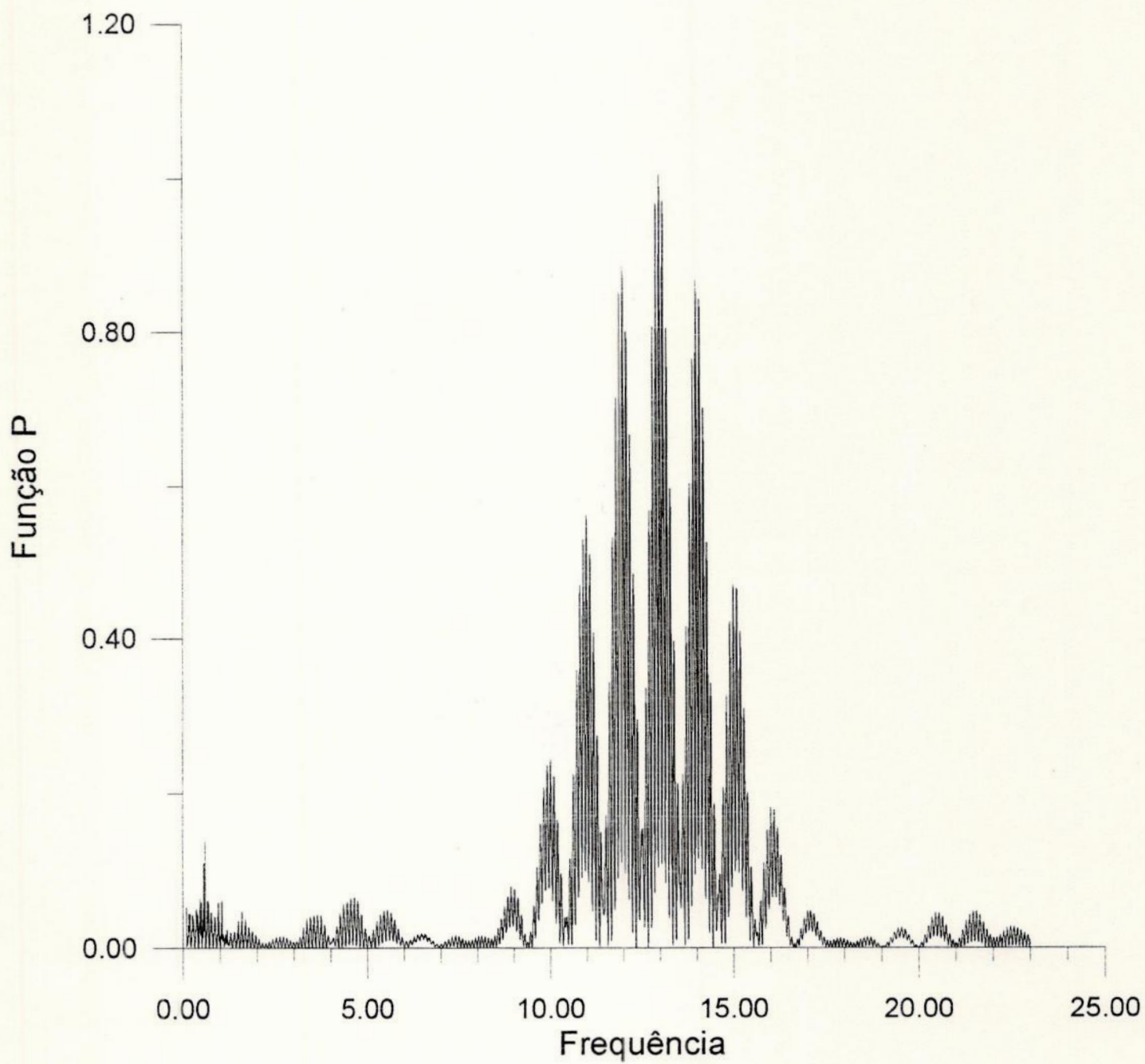


Fig 31- Periodograma Ponman para HD 93044 (Caso da Função Janela Senoidal)

```
program JURKEVICH;
```

```

-----
{ Este programa tem por objetivo determinar o periodo de series temporais }
{ desigualmente espacadas via metodo descrito pelo artigo "A METHOD OF }
{ COMPUTING PERIODS OF CYCLIC PHENOMENA" de I. JURKEVICH(1971). }
{ RESUMO DO METODO:Dada uma serie temporal,nos fazemos um "chute" de seu }
{ periodo e criamos um intervalo de periodos que contenha o periodo chutado. }
{ A seguir,nos pegamos os valores de "t"(tempo) da serie e calculamos para }
{ cada periodo as fases correspondentes aos "t"s.Por exemplo,se nos temos um }
{ intervalo de periodos que varie de 20 a 30 e uma serie com 50 pontos,tere- }
{ mos fases dadas por :fase[1]=t[1]/20;fase[2]=t[2]/20;...;fase[50]=t[50]/20 }
{ para o periodo 20.O mesmo procedimento se repetira para os outros perio- }
{ dos.Estas fases so poderao ter valores entre 0 e 1(obs.:quando o valor da }
{ fase atingir 1,ele passa automaticamente a valer 0 em virtude da funcao }
{ MOD).Calculadas as fases para o primeiro periodo tentativa(no caso de nos- }
{ so exemplo,o 20),teremos para cada ponto da serie uma fase(como se nos es- }
{ tivessesmos tentando ajustar uma curva de periodo 20 a serie e tendo dife- }
{ rencas de fase entre os pontos da serie e os pontos da curva do periodo }
{ tentativa 20).Cada magnitude da serie tera,portanto,uma fase associada . }
{ Pegaremos agora o intervalo entre 0 e 1(intervalo permitido de valores pa- }
{ ra as fases)e o dividiremos,de forma abstrata,em pequenas "caixas",isto e, }
{ se dividirmos o intervalo entre 0 e 1 em 5 caixas,a primeira caixa iria de }
{ 0 ao valor imediatamente inferior a 0.2 que estivesse contido no conjunto }
{ de fases calculadas;a segunda caixa iria de 0.2 ao valor imediatamente in- }
{ ferior a 0.4 que estivesse contido no conjunto de fases calculadas;ate ao }
{ ultimo intervalo que iria de 0.8 ao valor imediatamente inferior a 1 que }
{ estivesse contido no conjunto de fases calculadas.Teriamos,entao,5 caixas }
{ com essas fronteiras delimitadas e as 50 fases com suas respectivas magni- }
{ tudes da serie(no caso do nosso exemplo).A partir dai,nos "guardamos" em }
{ cada caixa as magnitudes cuja fase correspondente estivesse compreendida }
{ nas fronteiras de cada caixa.Temos,portanto,5 conjuntos de magnitudes de- }
{ terminadas pelas fases.Calculamos a variancia de cada conjunto(v1) e faze- }
{ mos um somatorio destas variancias,que sera a soma das variancias(ou des- }
{ vios medios quadraticos) das caixas(vm2).Repetimos todos estes calculos }
{ para os outros periodos tentativa(21 ate 30) e o periodo tentativa que ti- }
{ ver o menor vm2(dispersao menor entre a curva artificial do periodo tenta- }
{ tiva e a serie) sera o periodo da serie temporal calcluado pelo metodo. }
{ OBS.:Este programa pode refinar este periodo,ja que o intervalo de perio- }
{ dos e flexivel e determinado pelo usuario. }
-----

```

```
uses crt,printer;
```

```
var
```

```

  cis,ert,a,mi,ms,wet,qwe,m,x,t,ncx,ca,c,ax,q:integer;
  fm,li,ls,min,pmin,tempo,magnitude,varfora,perfora:real;
  m1,xm,s,me,wo,vl:array[1..20] of real;
  vm2,b,bartlet:array[1..1000]of real;
  pint:array[1..1000]of real;
  magin,fase,ii:array[1..1000] of real;
  mml,priparc,segparc:array[1..30] of real;
  saida,dados:text;
  entra,nome:string[20];
label 1,2;

```

```

begin
repeat
qwe:=0;
{-----bloco que le os dados da serie (arquivo)-----}
writeln('de o nome do arquivo(maximo de 20 caracteres)');
readln(entra);
assign(dados,entra);
reset(dados);
repeat
read(dados,tempo);
readln(dados,magnitude);
qwe:=qwe+1;
magin[qwe]:=magnitude;
ii[qwe]:=tempo;
writeln('magnitude ',magin[qwe]:3:3,' tempo ',ii[qwe]:3:3);
until eof(dados);
writeln('numero de dados do arquivo ',qwe:3);
{-----Leitura do numero de caixas e intervalo de periodos-----}
writeln('de o numero de caixas(maximo de 20)');
readln(ncx);
1:writeln('de os limites inferior e superior do intervalo');
writeln('para o periodo tentativa');
readln(li,ls);
writeln('de a ordem de grandeza das subdiviso'es');
writeln('do intervalo(1,10,100,1000,10000,etc.)');
readln(fm);
mi:=round(li*fm);
ms:=round(ls*fm);
{--Inicio do bloco principal de calculo do programa--}
for m:=mi to ms do
begin
for ca:=1 to ncx do
begin
ml[ca]:=0; xm[ca]:=0; {sub-bloco 1: }
s[ca]:=0; me[ca]:=0; {variaveis indexadas}
wo[ca]:=0; vl[ca]:=0; {auxiliares iniciais}
end;
{fim sub-bloco 1}
for t:= 1 to qwe do
begin
{sub-bloco 2: }
pint[m]:=(1/fm)*m; {calculo do periodo tentativa}
fase[t]:=frac(ii[t]/pint[m]);{calculo das fases}
for c:=1 to ncx do
begin
{sub-bloco 3(contido }
if fase[t]>=((c-1)/ncx) then {no sub-bloco 2): }
if fase[t]<(c/ncx) then {determinacao dos }
begin {intervalos das cai- }
ml[c]:=ml[c]+1; {xas e calculo de }
s[c]:=s[c]+magin[t]; {seus vl }
wo[c]:=wo[c]+sqr(magin[t]);
me[c]:=s[c]/ml[c];
xm[c]:=-ml[c]*sqr(me[c]);
vl[c]:=wo[c]-xm[c];

```

```

        end;
        end;{fim sub-bloco 3}
        end;{fim sub-bloco 2}
        for cis:=1 to ncx do
            begin
writeln(' valor de vl2 ',vl[cis]:3:9,'
valor de ml ',ml[cis]:3:3,' periodo ',pint[m]:3:6);
end;

        for ert:=1 to ncx do
            begin
                mml[ert]:=mml[ert-1]+((ml[ert]-1)*vl[ert]);
            end;
            priparc[ert]:=ln(mml[ncx]/(qwe-ncx));
            for ert:=1 to ncx do begin
                segparc[ert]:=segparc[ert-1]+((ml[ert]-1)*ln(vl[ert]));
            end;
            b[ert]:=((qwe-ncx)*priparc[ert])-segparc[ert];
            bartlet[m]:=b[ert];
            for ax:=1 to (ncx-1) do
                begin
                    vl[ax+1]:=vl[ax]+vl[ax+1];
                    vm2[m]:=vl[ax+1];
                end;
            end;
        end; {fim do bloco principal de calculos}
        for x:=mi to ms do
            begin
                {bloco de mostra de resultados na tela}
                writeln(' valor de vm2 e ',vm2[x]:6:6,
                    ' periodo ',pint[x]:6:6, ' bartlett ',bartlett[x]:3:9);
            end;
        min:=vm2[mi];
        for x:=mi to ms do
            begin
                {bloco de calculo do valor minimo}
                if vm2[x]<min then
                    begin
                        min:=vm2[x];
                        pmin:=x/fm;
                    end;
            end;
        writeln('valor minimo e ',min:3:3);
        writeln(' periodo correspondente e ',pmin:6:6);
        writeln('de um nome para o arquivo que vai conter estes dados');
        {----bloco de armazenamento dos resultados em arquivo----}
        readln(nome);
        assign(saida,nome);
        rewrite(saida);
        for wet:=mi to ms do
            begin
                varfora:=pint[wet];
                perfora:=vm2[wet];
                write(saida,varfora);
                writeln(saida,perfora);
            end;

```

```
close(saida);
{----artificio para refinar os resultados----}
writeln('Deseja obter um valor mais preciso para o periodo?(Sim=1,Nao=0)');
readln(a);
if a=1 then goto 1;
writeln('para voltar ao turbo pascal digite 0');
readln(q);
until q=0;
end.{fim do programa}
```

```
program PONMAN1;
{PROGRAMA QUE CALCULA OS NUMERADORES DE A E B}
uses crt,printer;
var mi,ms,zero,m,que,t,wet,wit:integer;name,nome,entra:string[20];
sai,saida,dados:text;
li,ls,magnitude,tempo,fm,varfora,perfora,freqout:real;
magin,ii,ai,bi:array [1..300] of real;
a,b,freq:array [1..2000] of real;
begin
repeat
que:=0;
writeln('De o nome do arquivo a ser lido (maximo de 20 caracteres)');
readln(entra);
assign(dados,entra);
reset(dados);
repeat
read(dados,tempo);
readln(dados,magnitude);
que:=que+1;
magin[que]:=magnitude;
ii[que]:=tempo;
writeln(' magnitude ',magin[que]:3:3,' tempo ',ii[que]:3:3);
until eof(dados);
writeln('numero de dados do arquivo ',que:3);
{_____}
writeln('Entre com os limites inferior e superior');
writeln('do intervalo de frequencias');
readln(li,ls);
writeln('De a ordem de grandeza das subdiviso'es do intervalo');
writeln('(1,10,100,1000,etc)');
readln(fm);
mi:=round(li*fm);
ms:=round(ls*fm);
for m:=mi to ms do
begin
for t:=1 to que do
begin
freq[m]:=m/(1*fm);
ai[t]:=ai[t-1]+(magin[t]*cos(2*pi*freq[m]*ii[t]));
bi[t]:=bi[t-1]+(magin[t]*sin(2*pi*freq[m]*ii[t]));
end;
a[m]:=ai[que];
b[m]:=bi[que];writeln('a ',a[m]:3:9,' b ',b[m]:3:9,
' freq ',freq[m]:3:9);
end;
writeln('De um nome para o arquivo que vai conter estes dados');
readln(nome);
assign(saida,nome);
rewrite(saida);
for wet:=mi to ms do
begin
varfora:=a[wet];
```

```
perfora:=b[wet];
write(saida,varfora,' ');
writeln(saida,perfora);
end;
close(saida);
writeln('de um nome para o arquivo que vai conter ');
writeln('os dados da frequencia');
readln(name);
assign(sai,name);
rewrite(sai);
for wit:=mi to ms do
begin
freqout:=freq[wit];
writeln(sai,freqout);
end;
close(sai);
writeln('para voltar ao turbo pascal digite 0');
readln(zero);
until zero=0;
end.
```

```
program ponman1a;
{PROGRAMA QUE CALCULA OS DENOMINADORES DE A E B}
uses printer;
var zero,que,wet,mi,ms,m,t:integer;dados,saida:text;nome,entra:string[20];
tempo,li,ls,fm,varfora,perfora:real;
ii,x,y:array [1..300] of real;
xis,yis,freq:array [1..2000] of real;
begin
repeat
que:=0;
writeln('De o nome do arquivo da serie temporal');
readln(entra);
assign(dados,entra);
reset(dados);
repeat
readln(dados,tempo);
que:=que+1;
ii[que]:=tempo;
writeln(' tempo ',ii[que]:3:3);
until eof(dados);
writeln('numero de dados do arquivo ',que:3);
{-----}
writeln('De os limites superior e inferior do intervalo de frequencias');
readln(li,ls);
writeln('de a ordem de grandeza das subdivisooes do intervalo');
writeln('mesma de ponman1');
readln(fm);
mi:=round(li*fm);
ms:=round(ls*fm);
writeln('Ponman2:calculo dos denominadores');
for m:=mi to ms do
begin
for t:=1 to que do
begin
freq[m]:=m/fm;
x[t]:=x[t-1]+sqr(cos(2*pi*ii[t]*freq[m]));
y[t]:=y[t-1]+sqr(sin(2*pi*ii[t]*freq[m]));
end;
xis[m]:=x[t];
yis[m]:=y[t];
writeln('valor xis ',xis[m]:3:9,
' valor yis ',yis[m]:3:9);
end;
writeln('De um nome ao arquivo que vai conter estes dados');
readln(nome);
assign(saida,nome);
rewrite(saida);
for wet:=mi to ms do
begin
varfora:=xis[wet];
perfora:=yis[wet];
write(saida,varfora,' ');
```



```
writeln(saida,perfora);  
end;  
close(saida);  
writeln('para sair do turbo pascal digite 0');  
readln(zero);  
until zero=0;  
end.
```

```
program ponman2;
{PROGRAMA QUE CALCULA A E B}
uses printer;
var
zero,que,quei,t,m,mi,ms,wet:integer;entre,entra,nome:string[20];
saida,dedos,dados:text;
magnitude,tempo,a,b,li,ls,fm,xo,yo,varfora,perfora:real;
xis,yis,ai,bi:array [1..2000] of real;
begin
repeat
quei:=0;
que:=0;
writeln('De o nome do arquivo dos denominadores');
readln(entre);
assign(dedos,entre);
reset(dedos);
repeat
read(dedos,xo);
readln(dedos,yo);
quei:=quei+1;
xis[quei]:=xo;
yis[quei]:=yo;
until eof(dedos);
writeln('numero de dados da serie ',quei:3);
writeln('De o nome do arquivo que contem a,b e frequencia');
readln(entra);
assign(dados,entra);
reset(dados);
repeat
read(dados,a);
readln(dados,b);
que:=que+1;
ai[que]:=a;
bi[que]:=b;
until eof(dados);
writeln('numero de dados do arquivo ',que:3);
writeln('Entre com os limites inferior e superior do intervalo');
writeln('de frequencias');
readln(li,ls);
writeln('De a ordem de grandeza das subdivisoes do intervalo');
writeln('(a mesma de PONMAN1)');
readln(fm);
mi:=round(li*fm);
ms:=round(ls*fm);
writeln('ponman 2:divisao entre numerador e denominador');
for t:=1 to quei do
begin
ai[t]:=ai[t]/xis[t];
bi[t]:=bi[t]/yis[t];
writeln('ai ',ai[t]:3:9,' bi ',bi[t]:3:9);
end;
writeln('de um nome ao arquivo que vai conter estes dados');
```

```
readln(nome);
assign(saida,nome);
rewrite(saida);
for wet:=1 to que do
begin
varfora:=ai[wet];
perfora:=bi[wet];
write(saida,varfora,' ');
writeln(saida,perfora);
end;
close(saida);
writeln('para voltar ao turbo pascal digite 0');
readln(zero);
until zero=0;
end.
```

```
program ponman3;
{PROGRAMA QUE CALCULA OS PONTOS DO PERIODOGRAMA}
var
zero,que,qui,t,wet,zerum:integer;nome,entra,entre:string[20];
dedos,saida,dados:text;
perfora,varfora,ampla,amplb,f:real;
p,amplita,amplitb,freq:array [1..2001] of real;
begin
repeat
que:=0;
qui:=0;
writeln('entre com o arquivo das frequencias');
readln(entre);
assign(dados,entre);
reset(dados);
repeat
readln(dados,f);
qui:=qui+1;
freq[qui]:=f;
writeln(' freq ',freq[qui]:3:9);
until eof(dados);
writeln('numero de dados do arquivo ',qui:3);
writeln('entre com o arquivo das amplitudes a e b');
readln(entra);
assign(dados,entra);
reset(dados);
repeat
read(dados,ampla);
readln(dados,amplb);
que:=que+1;
amplita[que]:=ampla;
amplitb[que]:=amplb;
writeln('amplitude a ',amplita[que]:3:3,' amplitude b ',amplitb[que]:3:3);
until eof(dados);
writeln('numero de dados do arquivo ',que:3);
writeln('para continuar digite 0');
readln(zerum);
for t:=1 to que do
begin
p[t]:=sqr(amplita[t])+sqr(amplitb[t]);
writeln(' valor de p ',p[t]:3:9);
end;
writeln('de um nome ao arquivo que vai conter estes dados');
readln(nome);
assign(saida,nome);
rewrite(saida);
for wet:=1 to que do
begin
varfora:=p[wet];
perfora:=freq[wet];
write(saida,perfora,' ');
writeln(saida,varfora);
```

```
end;  
close(saida);  
writeln('para voltar ao turbo pascal digite 0');  
readln(zero);  
until zero=0;  
end.
```

```
program ponman4;
{PROGRAMA QUE CALCULA A VARIANCIA DOS VALORES DE P}
uses printer;
var
quo, zero, que, qui, wet, m, mi, ms, t, qual, mix: integer;
dados, dados, dados, saida: text;
entrou, entra, entre, nome: string[20]; p, magnitude, tempo, sigma, li, ls, fm, varpop,
condicao, varfora, perfora, titi, f: real;
magin, x: array [1..300] of real;
pin, varian, freq, varien: array [1..2000] of real;
begin
repeat
que:=0;
writeln('De o nome do arquivo que contem p');
readln(entra);
assign(dados, entra);
reset(dados);
repeat
read(dados, titi);
readln(dados, p);
que:=que+1;
pin[que]:=p;
until eof(dados);
writeln('numero de dados do arquivo ', que:3);
qui:=0;
writeln('De o nome do arquivo da serie temporal');
readln(entre);
assign(dados, entre);
reset(dados);
repeat
read(dados, tempo);
readln(dados, magnitude);
qui:=qui+1;
magin[qui]:=magnitude;
until eof(dados);
writeln('numero de dados do arquivo ', qui:3);
quo:=0;
writeln('entre com o arquivo das frequencias');
readln(entrou);
assign(dados, entrou);
reset(dados);
repeat
readln(dados, f);
quo:=quo+1;
freq[quo]:=f;
until eof(dados);
writeln('numero de dados do arquivo ', quo:3);
writeln('de o valor do erro das medidas');
readln(sigma);
mi:=round(li*fm);
ms:=round(ls*fm);
writeln('Ponman4: calculo das variancias');
```

```
    for m:=1 to que do
    begin
    for t:=1 to qui do
    begin
    x[t]:=x[t-1]+(sqr(magin[t]));
    end;
    varian[m]:=(8*pin[m])/(qui/sigma);
    writeln('variancia ',varian[m]:3:13
    , 'freq ',freq[m]:3:3);
    end;
varpop:=x[t]/qui;
writeln('varpop ',varpop:3:3);
condicao:=8*varpop/sigma;
writeln('numero de dados da serie temporal ',qui:3);
writeln('valor do termo condicional ',condicao:3:3);
writeln('Deseja adicionar variancia devida a cancelamento imperfeito ');
writeln('de termos cruzados `a variancia dos erros ');
writeln('observacionais?(sim=1,nao=0)');
readln(qual);
if qual=1 then begin
for mix:=1 to qui do
begin
varien[mix]:=varian[mix]+((16/qui)*(varpop*pin[mix]));
writeln('ja que voce quis mudar,ai vai:variancia agora e ',varien[mix]:3:3);
end;
end;
writeln('de um nome ao arquivo que vai conter esses dados');
readln(nome);
assign(saida,nome);
rewrite(saida);
for wet:=1 to qui do
begin
varfora:=pin[wet];
if qual=1 then begin
perfora:=varien[wet];
end;
if qual<>1 then begin
perfora:=varian[wet];
end;
write(saida,varfora,' ');
writeln(saida,perfora);
end;
close(saida);
writeln('para voltar ao turbo pascal digite 0');
readln(zero);
until zero=0;
end.
```

```
program alias1;
{PROGRAMA PARA CALCULO DE NUMERADORES DE AG E BG}
uses printer;
var
zero,que,wat,wet,mi,ms,m,t:integer;
frequencia,tempo,varfora,perfora,freqfora,fm,li,ls,fase,tint:real;
entra,nome,name:string[20];dados,saida,sai:text;
x,y,ii:array [1..300] of real;
xi,yi,freq:array [1..2000] of real;
begin
repeat
que:=0;
writeln('de o nome do arquivo da serie temporal');
readln(entra);
assign(dados,entra);
reset(dados);
repeat
readln(dados,tempo);
que:=que+1;
ii[que]:=tempo;
until eof(dados);
writeln('numero de dados do arquivo ',que:3);
{-----}
writeln('de os limites inferior e superior do intervalo');
writeln('de frequencias');
readln(li,ls);
writeln('de a ordem de grandeza das subdivisoes(mesma de Ponman1)');
readln(fm);
mi:=round(li*fm);
ms:=round(ls*fm);
writeln('entre com o valor da fase');
readln(fase);
writeln('agora entre com o valor do tempo de integracao(em fracao de dia)');
readln(tint);
writeln('entre com o valor da frequencia do sinal senoidal');
readln(frequencia);
writeln('alias1:calculo dos numeradores');
for m:=mi to ms do
begin
for t:=1 to que do
begin
freq[m]:=m/fm;
x[t]:=x[t-1]+(cos(2*pi*frequencia*ii[t]+fase)*
cos(2*pi*freq[m]*ii[t])*
sin(frequencia*tint*pi)/(frequencia*tint*pi));
y[t]:=y[t-1]+(cos(2*pi*frequencia*ii[t]+fase)*
(sin(2*pi*freq[m]*ii[t]))*
sin(frequencia*tint*pi)/(frequencia*tint*pi));
end;
xi[m]:=x[t];
yi[m]:=y[t];
writeln(' ',xi[m]:3:9,' ',yi[m]:3:9,freq[m]:2:3);
```



```
end;
writeln('de o nome do arquivo que vai conter esses dados');
readln(nome);
assign(saida,nome);
rewrite(saida);
  for wet:=mi to ms do
  begin
    varfora:=xi[wet];
    perfora:=yi[wet];
    write(saida,varfora,' ');
    writeln(saida,perfora);
  end;
close(saida);
writeln('de um nome para o arquivo que vai conter os dados ');
writeln('do intervalo de frequencias');
readln(name);
assign(sai,name);
rewrite(sai);
  for wat:=mi to ms do
  begin
    freqfora:=freq[wat];
    writeln(sai,freqfora);
  end;
close(sai);
writeln('para voltar ao turbo pascal digite 0');
readln(zero);
until zero=0;
end.
```

```
program alias2;
{PROGRAMA QUE CALCULA AG E BG}
uses printer;
var
zero,que,qui,t,wet:integer;
dados,dedos,saida:text;entre,entra,nome:string[20];
xi,yi,c,s,varfora,perfora,fm:real;
nc,ns,dc,ds:array [1..2000] of real;
begin
repeat
que:=0;
qui:=0;
writeln('de o nome do arquivo que contem o numerador(para alias)');
readln(entra);
assign(dados,entra);
reset(dados);
repeat
read(dados,xi);
readln(dados,yi);
que:=que+1;
nc[que]:=xi;
ns[que]:=yi;
until eof(dados);
writeln('numero de dados do arquivo ',que:3);
writeln('de o nome do arquivo que contem os termos em denominador');
readln(entre);
assign(dedos,entre);
reset(dedos);
repeat
read(dedos,c);
readln(dedos,s);
qui:=qui+1;
dc[qui]:=c;
ds[qui]:=s;
until eof(dedos);
writeln('numero de dados do arquivo ',qui:3);
{-----}
writeln('digite 0 para continuar');
readln(fm);
writeln('alias2:calculo das amplitudes ag e bg');
for t:=1 to que do
begin
nc[t]:=nc[t]/dc[t];
ns[t]:=ns[t]/ds[t];
writeln(nc[t]:3:9,' ',ns[t]:3:9);
end;
writeln('de um nome ao arquivo que vai conter esses dados');
readln(nome);
assign(saida,nome);
rewrite(saida);
for wet:=1 to qui do
begin
```

```
varfora:=nc[wet];
perfora:=ns[wet];
write(saida,varfora,' ');
writeln(saida,perfora);
end;
close(saida);
writeln('para voltar ao turbo pascal digite 0');
readln(zero);
until zero=0;
end.
```

```
program alias3;
{PROGRAMA QUE GERA OS PONTOS DO PERIODOGRAMA DA FUNCAO SENOIDAL}
uses printer;
var
qui, zero, que, m, t, wet, fm: integer;
ampla, amplb, varfora, perfora, fr: real;
ag, bg, p, pg, freq: array [1..2000] of real;
dados, dedos, saida: text; entra, nome, entre: string[20];
begin
repeat
que:=0;
qui:=0;
writeln('entre com o nome do arquivo que contem ag e bg(caso alias)');
readln(entra);
assign(dados, entra);
reset(dados);
repeat
read(dados, ampla);
readln(dados, amplb);
que:=que+1;
ag[que]:=ampla;
bg[que]:=amplb;
until eof(dados);
writeln('numero de dados do arquivo ', que:3);
writeln('de o nome do arquivo que contem o intervalo de frequencias');
readln(entre);
assign(dados, entre);
reset(dados);
repeat
readln(dados, fr);
qui:=qui+1;
freq[qui]:=fr;
until eof(dados);
{-----}
writeln('para continuar execucao digite 0');
readln(fm);
writeln('alias3: calculo de pg');
for t:=1 to que do
begin
p[t]:=sqr(ag[t])+sqr(bg[t]);
writeln('p ', p[t]:3:9);
end;
writeln('de um nome ao arquivo que vai conter esses dados');
readln(nome);
assign(saida, nome);
rewrite(saida);
for wet:=1 to que do
begin
varfora:=p[wet];
perfora:=freq[wet];
write(saida, perfora, ' ');
writeln(saida, varfora);
```

```
end;  
close(saida);  
writeln('para voltar ao turbo pascal digite 0');  
readln(zero);  
until zero=0;  
end.
```

```
program saimedia;
{PROGRAMA QUE SUBTRAI A MEDIA DOS DADOS DA SERIE TEMPORAL}
var
zero,t,wet,w,que,tau:integer;
magnitude,tempo,magfora,timefora,media:real;
dados,saida:text;nome,entra:string[20];
magin,ii,x,y:array [1..300] of real;
begin
repeat
que:=0;
writeln('de o nome do arquivo da serie temporal');
readln(entra);
assign(dados,entra);
reset(dados);
repeat
read(dados,tempo);
readln(dados,magnitude);
que:=que+1;
magin[que]:=magnitude;
ii[que]:=tempo;
writeln(' magnitude ',magin[que]:3:3,' tempo ',ii[que]:3:3);
until eof(dados);
writeln('numero de dados do arquivo ',que:3);
writeln('digite 0 para continuar');
readln(w);
for t:=1 to que do
begin
x[t]:=x[t-1]+magin[t];
end;
media:=x[que]/que;
for tau:=1 to que do
begin
y[tau]:=magin[tau]-media;
writeln(' serie subtraida da media ',y[tau]:3:3);
end;
writeln('de o nome da serie subtraida da media');
readln(nome);
assign(saida,nome);
rewrite(saida);
for wet:=1 to que do
begin
magfora:=y[wet];
timefora:=ii[wet];
write(saida,timefora,' ');
writeln(saida,magfora);
end;
close(saida);
writeln('para voltar ao Turbo Pascal digite 0');
readln(zero);
until zero=0;
end.
```

```
program filtrodif;
{PROGRAMA QUE APLICA O FILTRO DE DIFERENCAS}
{ALEM DE CALCULAR 1/T E A FREQUENCIA NYQUIST}
var
zero,t,que,wet,w:integer;
magnitude,tempo,magfora,timefora,te,nyquist,deltat,unosobret:real;
entra,nome:string[20];saida,dados:text;
ii,magin,x,y:array[1..300] of real;
begin
repeat
que:=0;
writeln('de o nome do arquivo que contem');
writeln('a serie temporal subtraida da media');
readln(entra);
assign(dados,entra);
reset(dados);
repeat
read(dados,tempo);
readln(dados,magnitude);
que:=que+1;
magin[que]:=magnitude;
ii[que]:=tempo;
writeln(' magnitude ',magin[que]:3:3,' tempo ',ii[que]:3:3);
until eof(dados);
writeln('numero de dados do arquivo ',que:3);
writeln('digite 0 para continuar execucao');
readln(w);
for t:=2 to que do
begin
x[t]:=magin[t]-magin[t-1];
y[t]:=ii[t]-ii[t-1];
writeln('filtro de diferencas:novo valor da magnitude e ',x[t]:3:3
,' deltat ',y[t]:3:3);
end;
for t:=2 to que do
begin
y[t]:=y[t-1]+y[t];
end;
deltat:=(2*y[t])/(que-1);
nyquist:=1/deltat;
te:=ii[que]-ii[1];
unosobret:=1/te;
writeln('freq Nyquist ',nyquist:3:6,' 1/T ',unosobret:3:6);
writeln('de um nome ao arquivo que vai conter os dados');
writeln('da serie do filtro de diferencas');
readln(nome);
assign(saida,nome);
rewrite(saida);
for wet:=1 to que do
begin
magfora:=x[wet];
timefora:=ii[wet];
```

```
write(saida,timefora);
writeln(saida,magfora);
end;
close(saida);
writeln('para voltar ao turbo pascal digite 0');
readln(zero);
until zero=0;
end.
```