

Carlos A. da Silva Felgueiras

Caracterização Fractal de
S. de Batimentos C. Fetais

DLEC
FEUP
1996

Carlos Alexandre da Silva Felgueiras

**Caracterização Fractal de Sequências de Batimentos
Cardíacos Fetais**

0437

Cat
681.31(m)/ FELC/CAR

UNIVERSIDADE DO PORTO
Faculdade de Engenharia
BIBLIOTECA
N.º 6373
CDU
Data

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

SETEMBRO DE 1996

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Departamento de Engenharia Electrotécnica e de Computadores

Caracterização fractal de sequências de batimentos cardíacos fetais

Carlos Alexandre da Silva Felgueiras

Licenciado em Física/Matemática Aplicada
pela
Faculdade de Ciências da Universidade do Porto

Dissertação submetida para satisfação parcial dos
requisitos do grau de mestre
em
Engenharia Electrotécnica e de Computadores
(Área de especialização de Informática Industrial)

Dissertação realizada sob a supervisão de
Professor Doutor Joaquim Pontes Marques de Sá
do Departamento de Engenharia Electrotécnica e de Computadores
da Faculdade de Engenharia da Universidade do Porto

Porto, Setembro de 1996

DEEC

FEUP

*a Os Todos
e aos outros*

Agradecimentos

Ao Professor Marques de Sá pela paciência e disponibilidade com que orientou a presente dissertação.

Ao Doutor João Bernardes por facultar os dados e respectivos dados clínicos, que permitiram a realização deste trabalho, e a orientação dada em relação a alguns aspectos médicos do problema estudado.

Ao Professor Sílvio Gama, do Departamento de Engenharia Civil, pelas discussões genéricas sobre fractais que permitiram compreender os problemas da abordagem escolhida.

À JNICT pelo apoio financeiro (bolsa PRAXIS XXI/BM/2340/94) que permitiu a realização da dissertação a tempo inteiro.

Ao INEB pelas condições de trabalho disponibilizadas.

Agradeço a forma como tornaram possível a realização desta dissertação.

1 INTRODUÇÃO.....	1
1.1 OBJECTIVOS.....	1
1.2 MÉTODOS E MOTIVAÇÃO.....	1
1.3 FERRAMENTAS DE DESENVOLVIMENTO	2
1.4 CONTRIBUIÇÕES.....	3
1.5 ESTRUTURA.....	3
2 A SEQUÊNCIA DE BATIMENTOS CARDÍACOS FETAIS (FHR)	5
2.1 DEFINIÇÃO	5
2.2 CLASSES.....	7
2.3 MÉTODO DE RECONSTRUÇÃO DE SEQUÊNCIAS FHR.....	12
3 FRACTAIS.....	21
3.1 ALGUMAS DEFINIÇÕES.....	21
3.2 MODELOS FRACTAIS.....	26
3.2.1 Movimento browniano.....	26
3.2.2 Movimentos brownianos fraccionários (fBm).....	31
3.2.3 Taxinomia	34
3.3 GERAÇÃO DE FORMAS FRACTAIS	34
3.3.1 Geração de fractais geométricos.....	35
3.3.2 Geração de fractais estatísticos.....	35
3.3.2.1 Movimentos brownianos fraccionários (fractais espaciais).....	35
3.3.2.2 Funções de Brown fraccionárias (fractais temporais).....	36
3.3.2.2.1 Método do deslocamento do ponto médio.....	36
3.3.2.2.2 Método das adições sucessivas	38
3.3.2.2.3 Método da síntese espectral	39
3.4 ESTIMAÇÃO DE DIMENSÕES FRACTAIS.....	42
3.4.1 Dimensão box-counting.....	42
3.4.2 Dimensões fractais de fractais temporais.....	45
3.4.2.1 Expoente espectral.....	47
3.4.2.2 Método de Hugson (CGSA).....	49
3.4.2.3 Método de Higushi	55
4 TEORIA DO CAOS.....	59
4.1 ALGUMAS DEFINIÇÕES.....	59
4.2 DIMENSÃO FRACTAL DE ATRACTORES ESTRANHOS	61
4.2.1 Dimensão de informação e dimensão de correlação.....	61
4.2.2 Dimensão de Procaccia-Grassberger	63
5 CARACTERIZAÇÃO FRACTAL DE SEQUÊNCIAS FHR.....	67
5.1 CARACTERIZAÇÃO FRACTAL DOS TRAÇADOS DA SEQUÊNCIA FHR.....	67
5.2 CARACTERIZAÇÃO FRACTAL DA SEQUÊNCIA FHR NO ESPAÇO DE FASE.....	72
5.3 ANÁLISE DE RESULTADOS	75
6 CONCLUSÕES.....	83
7 REFERÊNCIAS BIBLIOGRÁFICAS	85
8 APÊNDICES	87
8.1 PROGRAMAS DE RECONSTRUÇÃO DA SEQUÊNCIA FHR	87
8.2 PROGRAMAS DE GERAÇÃO DE FORMAS FRACTAIS.....	92
8.3 PROGRAMAS DE DETERMINAÇÃO DE DIMENSÕES FRACTAIS	94
8.4 PUBLICAÇÕES	100

1 Introdução

1.1 Objectivos

Os sinais cardiotocográficos (CTG) são utilizados no diagnóstico do bem-estar do feto. O sinal cardiotocográfico é composto por três componentes: sequência de batimentos cardíacos fetais (FHR), pressão uterina (UP) e movimentos fetais (FM). A análise destes sinais é realizada em primeira instância por inspecção visual dos respectivos traçados em papel, por pessoal especializado. No entanto esta tarefa é de baixa produtividade e as características do CTG determinadas desta forma apresentam grande variabilidade, quer inter quer intra-observador.

O desenvolvimento de sistemas de aquisição digital de sinal, permitiu desenvolver diversos algoritmos computacionais que sistematizam algumas das operações efectuadas na inspecção visual dos traçados. Estes algoritmos baseiam-se, essencialmente, na análise de séries temporais, procurando determinar características estatísticas relevantes: média, moda, variância, desvios relativamente a um valor fixo e respectiva duração temporal, amplitudes dos sinais em escalas de tempo distintas.

Apesar de já existirem algumas aplicações informáticas que realizam a análise de sinais cardiotocográficos de forma muito satisfatória, por exemplo o SisPorto desenvolvido no INEB, o objectivo do presente trabalho é estudar a aplicação de modelos fractais para caracterização da sequência FHR. Esta abordagem é justificada por alguns resultados já publicados e, mesmo pela análise clássica de séries temporais, em especial, a sua caracterização em escala de tempo distintas.

1.2 Métodos e Motivação

Apesar de já terem sido inventados no início deste século, como contra-exemplos no estudo de funções, os objectos fractais só ganharam notoriedade após a divulgação da geometria fractal por Mandelbrot, nas décadas de 70-80. O interesse que despertou, criou um vasto leque de aplicações em diversas áreas. Entre elas podem citar-se:

Tabela 1-1: Aplicações dos fractais

Arte	Utilização de formas “abstractas”, como os conjuntos de Mandelbrot, com sentido estético, e de modelos fractais de paisagens.
Análise de séries temporais	Modelização de séries temporais de origem biológica, hidrológica, económica, etc.
Biologia	Além de modelos fractais para as séries temporais, também se modelizam estruturas espaciais, por exemplo, árvores ou plantas.
Electrónica	Modelização de ruídos do tipo 1/f característicos de muitas componentes electrónicas ou micro-electrónicas.
Física	Aplicações em física estatística e física da matéria condensada, em equivalência com os grupos de renormalização.
Hidrologia e Geologia	Aplicação de modelos fractais para lençóis freáticos (de superfície ou subterrâneos) ou jazigos minerais.
Matemática e Teoria do Caos	Descrição fractal de atractores estranhos de sistemas caóticos.
Processamento de Imagem	Técnicas de compressão de imagem.

Na área de análise de sinais biomédicos, as aplicações a sinais bidimensionais (imagens) consistem, essencialmente, na caracterização fractal de texturas como técnica de segmentação. Fortin 1992 aplicou estas técnicas a imagens por ressonância magnética do coração e a imagens de raios-X, com resultados satisfatórios, fundamentalmente, porque verificou serem insensíveis às variações de luminosidade e contraste presente neste tipo de imagens.

No domínio das séries temporais os objectos fractais têm sido aplicadas a sinais como o electroencefalograma, o electrocardiograma e a frequência de batimento cardíaco.

Relativamente a este último, a sua natureza fractal foi identificada por Goldberger 1992, baseando-se no espectro do tipo $1/f$ que esse sinal exhibe, e que já havia sido referido por Kobayashi 1982. Yamamoto 1993, usando outro método, baseado igualmente na análise espectral, confirma este resultado, comunicando que aproximadamente 85% da potência de sinal será devida a uma componente fractal. Baseado nestes resultados, Goldberger pôs em causa a teoria de homeostasia proposta por Walter B. Cannon, sugerindo que a variabilidade errática observada na sequência de batimentos cardíacos terá como causa os próprios sistemas fisiológicos e não causas externas a esses. Assim, a normalidade será observada como um estado aparentemente errático, enquanto estados estacionários (constantes ou periódicos) observados na sequência de batimentos serão indicadores de situações patológicas.

Quanto às sequências de batimentos cardíacos fetais têm sido divulgados resultados preliminares sobre as características fractais dos atractores construídos pela representação num espaço de fase da própria sequência, conforme as técnicas de análise propostas pela teoria do caos (Chaffin 1991, Schulman 1994).

Assim, no presente trabalho procurar-se-á estender os resultados obtidos por Goldberger à sequência de batimentos cardíacos, e numa segunda fase verificar se as técnicas da teoria do caos permitem obter uma caracterização útil deste tipo de sinais.

1.3 Ferramentas de Desenvolvimento

As ferramentas de desenvolvimento de software utilizadas foram o Matlab 4.0 e a linguagem C no ambiente do Microsoft Visual C++ 1.51. O Matlab permitiu a realização de um elevado número de experiências, aproveitando vários algoritmos da linguagem e das *toolboxes* especializadas, em particular a Signal Processing Toolbox, e as facilidades de visualização de resultados, tendo sido desenvolvida uma interface gráfica para realização de acções como: zoom em gráficos bidimensionais, mudanças de escala e regressão linear. Tratando-se de uma linguagem optimizada para operações vectoriais, a sua utilização nalguns dos algoritmos de análise propostos pela Teoria do Caos torna-se pouco eficiente, pelo que se optou por implementar esses algoritmos em linguagem C. Estes programas são inacessíveis a partir do ambiente Matlab, apesar de ser possível criar programas MEX executáveis nesse ambiente. A interface com o Matlab, onde é efectuada a visualização dos resultados, facilitada por uma interface gráfica específica, é feita por intermédio de ficheiros.

A análise estatística dos resultados foi efectuada com o pacote estatístico: SPSS for Windows 6.1.

1.4 Contribuições

No presente trabalho foi desenvolvido um novo método de reconstrução da sequência de batimentos cardíacos fetais (FHR), de forma a melhorar a qualidade da sequência obtida por monitores fetais convencionais.

Foi desenvolvida uma estrutura de dados em árvore para localização de pontos em espaços de dimensão superior a dois, que permite efectuar de forma eficiente o cálculo das dimensões fractais de um conjunto de pontos.

Relativamente aos resultados foi estabelecida a equivalência entre algumas das técnicas usadas trabalho e técnicas clássicas de análise. Verificando-se que as medidas mais conhecidas da teoria do caos (dimensão de atractores) não trazem mais valias para a análise e classificação das sequências FHR, sendo bem mais importante o volume ocupado pelos mesmos.

1.5 Estrutura

A presente dissertação está organizada em cinco capítulos. O primeiro constitui a introdução que termina nesta secção. O segundo introduz os sinais estudados, as técnicas de classificação actuais e a descrição do algoritmo de eliminação de artefactos da sequência FHR, desenvolvido no presente trabalho. O terceiro capítulo será dedicado aos fractais, focando as definições, modelos e estimadores de dimensões fractais. No quarto capítulo, introduz-se a teoria do caos e as técnicas de análise de séries temporais que utiliza. No quinto capítulo descreve-se a aplicação das técnicas expostas nos capítulos anteriores à caracterização das sequências de batimentos cardíacos fetais e os resultados obtidos, e no sexto as conclusões deste trabalho.

O sétimo capítulo lista as referências bibliográficas feitas ao longo do texto e o oitavo é reservado aos apêndices, com listagens dos programas mais relevantes e lista das publicações realizadas durante esta fase do curso de mestrado.

2 A sequência de batimentos cardíacos fetais (FHR)

Este capítulo é dedicado a uma descrição dos sinais cardiotocográficos e, em particular de uma das componentes, a sequência de batimentos cardíacos fetais (FHR- Fetal Heart Rate).

Na secção 2.1 define-se o sinal cardiotocográfico e as respectivas componentes, referem-se algumas das técnicas de classificação automática destes, assim como as medidas a que recorrem. Na secção 2.2 descrevem-se, com exemplos de cada caso, as classes de sequências FHR discriminadas por inspecção visual. Na secção 2.3 descreve-se o pré-processamento aplicado à sequência FHR com o objectivo de melhorar a sua qualidade, antes de se efectuar a sua análise.

2.1 Definição

A análise de sinais cardiotocográficos (CTG) é uma ferramenta muito utilizada no diagnóstico do bem-estar do feto. Este sinal compreende 3 componentes:

- Sequência de batimentos cardíacos fetais (FHR).
- Pressão uterina (UP), para detecção das contracções uterinas (UC).
- Movimentos fetais (FM), detectados pela paciente ou enfermeira.

O CTG é obtido através de monitores fetais, disponíveis nos serviços de Obstetrícia, que fornecem ao médico um traçado gráfico em papel contendo os sinais supracitados. Noutras variantes como no sistema SisPorto, desenvolvido no INEB, também é possível obter o CTG em formato digital.

Geralmente, a duração considerada mínima de um traçado é de vinte minutos, mas se a sua duração for superior o traçado pode ser dividido em secções com aquela duração. As técnicas de análise propostas neste trabalho são aplicadas a estas unidades, sem que se efectue mais nenhuma segmentação ou união das mesmas.

A inspecção visual do CTG pelo especialista médico permite-lhe obter informações sobre o estado clínico do feto. No entanto, este trabalho é bastante fastidioso e de baixa produtividade. Além disso, os resultados são bastante subjectivos, verificando-se uma elevada variabilidade inter-observador, e pouco reproduzíveis, com variabilidade intra-observador.

Por isso, tem sido crescente o interesse por métodos computacionais de classificação dos sinais cardiotocográficos. Algumas destas técnicas utilizam simultaneamente a sequência FHR e a pressão uterina (Sande 1981), procurando as relações entre ambas. Esta abordagem é corroborada por modelos de mecanismos fisiológicos responsáveis pelas formas da sequência FHR (Carrillo 1983). Outras técnicas propostas, efectuem a análise das três componentes de forma mais independente, apesar de alguns parâmetros para a sequência FHR serem afectados pelas outras componentes (Marques de Sá 1992). A correlação entre as três componentes é feita apenas sobre as características determinadas para cada uma delas (Dawes 1981), por exemplo, acelerações e ocorrência de movimentos fetais ou desacelerações e ocorrência de contracções uterinas. No presente trabalho procura-se caracterizar apenas a sequência de batimentos cardíacos fetais.

Esta sequência é, por definição, a série da frequência instantânea de batimento, ou seja, o inverso do período de batimento. Na Figura 2-1, representam-se os batimentos pelas setas, define-se o período de batimento T e a respectiva sequência de batimentos cardíacos pelos quadrados a cheio.

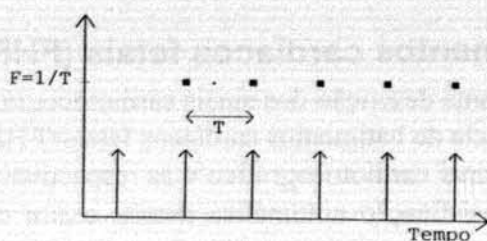


Figura 2-1: Sequência FHR constante

Assim, a variável medida pelos monitores fetais é o período de batimento, apesar de retornarem a frequência de batimento em batimentos por minuto (bpm). A medição daqueles intervalos temporais é realizada recorrendo a técnicas não invasivas, sendo as baseadas em ultrassons de Doppler, que detectam movimentos no coração fetal, as mais utilizadas. Outras técnicas utilizam os potenciais eléctricos detectados no ventre materno para estimar o ECG do feto, após a compensação do ECG materno (Dawes 1981), determinando o período de batimento a partir dos intervalos R-R do ECG fetal estimado.

As primeiras técnicas utilizando ultrassons de Doppler identificavam os batimentos como os máximos do sinal de Doppler superiores a um dado limiar (Tuck 1982). No entanto, esta técnica originava sequências com elevada incidência, em média 40% do tempo total da sequência, de amostras incorrectas (Dawes 1981). Tuck propôs e implementou um algoritmo que aplica técnicas de autocorrelação para procurar uma melhor estimativa do período de batimento. Esta melhoria advém do facto de utilizar mais informação contida na janela do sinal de Doppler, associada a cada batimento, do que um simples máximo, como na técnica anterior. Aquela informação extra é associada aos múltiplos eventos do ciclo cardíaco: abertura e fecho de válvulas, movimentos dos músculos e pulsação do sangue. Uma função de autocorrelação aplicada a janelas moventes, contendo 3 ou 4 batimentos, permite obter uma estimativa mais fiável do período, a partir do máximo dessa função e, portanto, diminuir significativamente a incidência de amostras erradas (artefactos). As sequências FHR utilizadas neste estudo, foram adquiridas com monitores fetais que utilizam técnicas de autocorrelação. A sua qualidade será discutida numa secção posterior.

A sequência FHR é, portanto, uma série temporal amostrada a uma frequência variável, dependente da própria série. Alguns autores (Dawes 1981, Dawes 1990) utilizam uma série temporal auxiliar, amostrada a frequência constante, obtida pela média das amostras originais em épocas com duração igual a 1/16 minuto. No entanto, neste trabalho e no seguimento dos trabalhos apresentados em Marques de Sá 1992, a análise é efectuada directamente sobre a sequência original. Portanto, na análise numérica da sequência FHR, considera-se que esta é uma série temporal, no sentido estrito, sobre a qual se pode, por exemplo, aplicar técnicas de análise espectral baseadas na Fast Fourier Transform (FFT).

A análise das sequências FHR, efectuada em Marques de Sá 1982, é orientada pelas características utilizadas na análise por inspecção visual das mesmas. São elas:

- **Linha base:** Frequência de batimento mais frequente na ausência de movimentos fetais e/ou contracções uterinas, determinada pelo máximo do histograma desses valores.
- **Acelerações e Desacelerações:** Conjunto de amostras da sequência que se afastam da linha base por períodos de tempo superiores a 15 segundos no sentido de valores crescentes ou decrescentes, respectivamente, de pelo menos 15 batimentos por minuto.
- **Variabilidade a longo e curto prazo:** Diferença entre o valor máximo e mínimo, amplitude (range), da sequência FHR em intervalos de tempo com duração de meio minuto e de dois segundos, respectivamente. A variabilidade a curto ou longo prazo pode caracterizar-se apenas pela sua ausência ou presença, conforme a respectiva amplitude é inferior ou superior a um dado limiar, dependente do tipo de variabilidade, ou pela sua taxa de incidência, a percentagem de tempo em que a variabilidade é superior àquele limiar.

2.2 Classes

Neste capítulo descrevem-se, de forma sucinta, doze classes de sequências FHR identificadas por inspecção visual do respectivo traçado, referindo o estado de saúde do feto que lhes está associado. Também se indicam, para cada classe, os valores típicos das características definidas anteriormente.

Salvo indicação em contrário a linha base dos traçados FHR, para as diversas classes, situa-se no intervalo 110-150 batimentos por minuto.

- **Classe A:** Classe de sequências FHR correspondente a uma situação fisiológica normal, com o feto num estado de sono profundo ou sono NEM (Non Eye Movements). É caracterizada pela ausência de acelerações e desacelerações, presença de variabilidade a curto prazo e variabilidade a longo prazo entre 5 a 10 batimentos por minuto. Exemplo na Figura 2-2.
- **Classe B:** Esta classe de sequências FHR também é indicadora de uma situação fisiológica normal, correspondente à fase de sono REM (Rapid Eye Movements). É caracterizada pela presença de acelerações e ausência de desacelerações, presença de variabilidade a curto prazo e a longo prazo apresenta valores superiores a 10 batimentos por minuto. Exemplo na Figura 2-3.

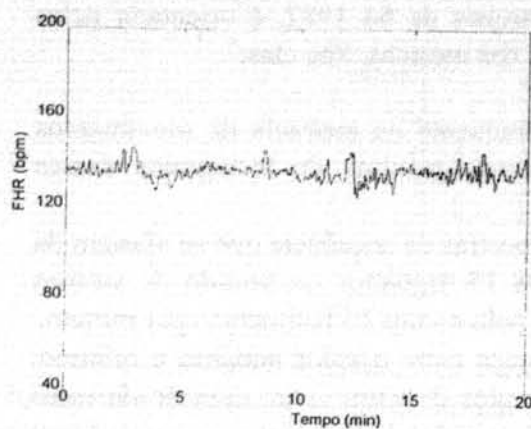


Figura 2-2: Sequência FHR da classe A

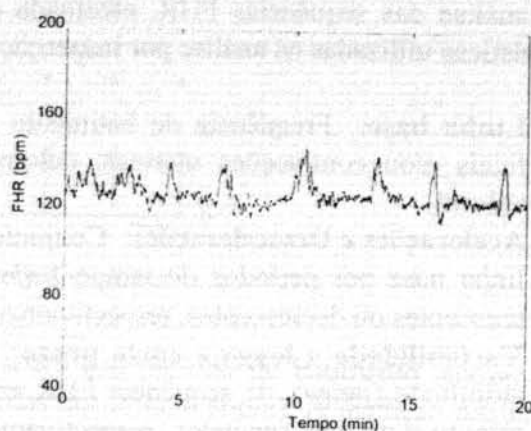


Figura 2-3: Sequência FHR da classe B

• **Classe C:** Outra classe normal, correspondente ao estado de vigília calma. As características desta classe são idênticas às da classe A, à exceção do valor da variabilidade a longo prazo, que para sequências da classe C é superior a 10 batimentos por minuto. Exemplo na Figura 2-4.

• **Classe D:** Continua a corresponder a uma situação fisiológica normal do feto, indicando que este se encontra numa situação de vigília activa. As características desta classe são idênticas às da classe B, distinguindo-se apenas pela maior amplitude e duração temporal das acelerações em sequências desta classe. Exemplo na Figura 2-5.

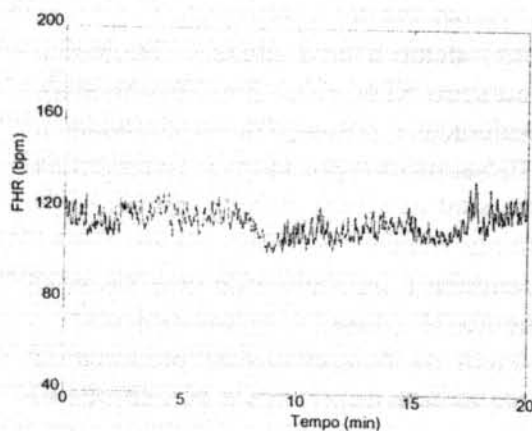


Figura 2-4: Sequência FHR da classe C

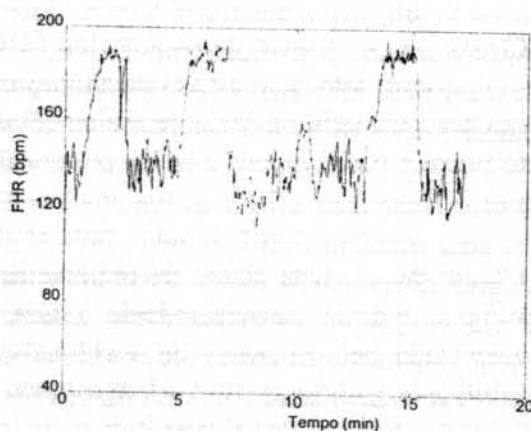


Figura 2-5: Sequência FHR da classe D

• **Classe AD:** *Accelerative and Decelerative*, esta classe de sequências FHR é indicadora da situação fisiológica anormal em que o feto se encontra, correspondendo a um estado de stress que ocorre, por exemplo, durante o parto. Tal como o nome indica, é caracterizada pela presença de acelerações e desacelerações, assim como variabilidade a curto prazo. A variabilidade a longo prazo é superior a 10 batimentos por minuto. Exemplo na Figura 2-6.

- **Classe DE:** *DEcelerative*, as sequências FHR desta classe indicam que o feto se encontra num estado de grande stress que pode ocorrer nas fases terminais do parto. Sequências desta classe são caracterizados por não apresentarem acelerações, verificando-se a presença de desacelerações. Podem ou não apresentar variabilidade a curto prazo e a longo prazo tem valores entre 5 e 10 batimentos por minuto. Exemplo na Figura 2-7.

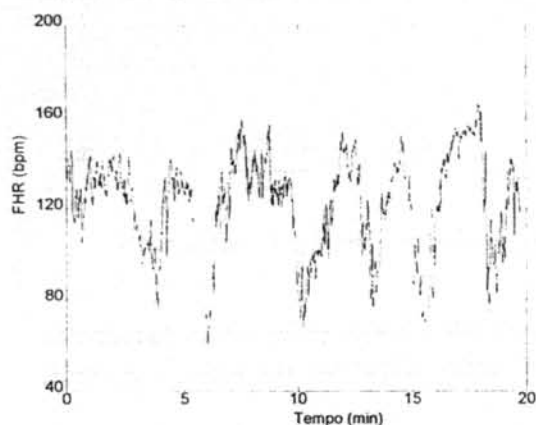


Figura 2-6: Sequência FHR da classe AD

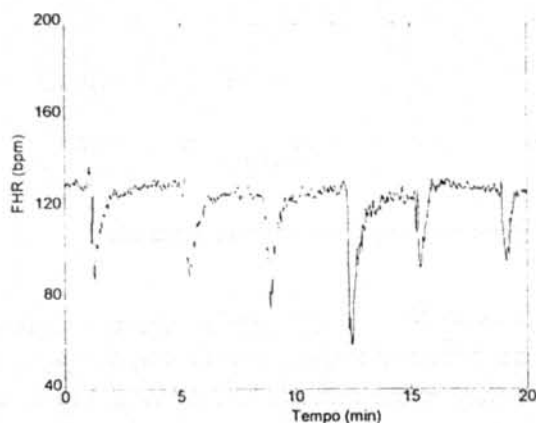


Figura 2-7: Sequência FHR da classe DE

- **Classe LD:** *Largely Decelerative*, esta classe distingue-se da classe DE apenas pela maior duração temporal das desacelerações, em geral superiores a dois minutos para sequências pertencentes a esta classe. Exemplo na Figura 2-8.

- **Classe FS:** *Flat Sinusoidal*, as sequências FHR pertencentes a esta classe indicam uma situação fisiológica anormal, com o feto numa situação de sofrimento crónico, devido, por exemplo, a insuficiências placentárias. São características das sequências desta classe a ausência de acelerações, desacelerações e variabilidade a curto prazo (em pelo menos 50% do traçado). A variabilidade a longo prazo apresenta, também, valores muito baixos, inferiores a 5 batimentos por minuto. Por vezes podem observar-se algumas desacelerações. Exemplo na Figura 2-9.

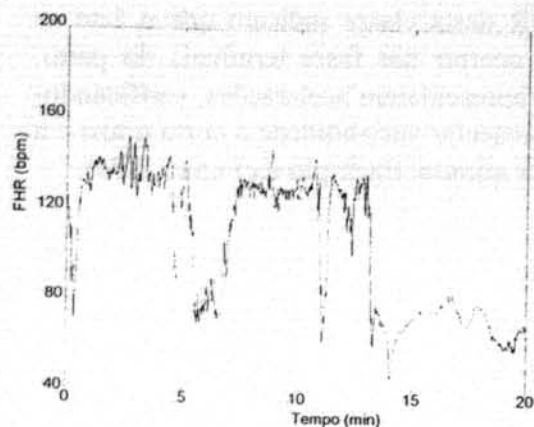


Figura 2-8: Sequência FHR da classe LD

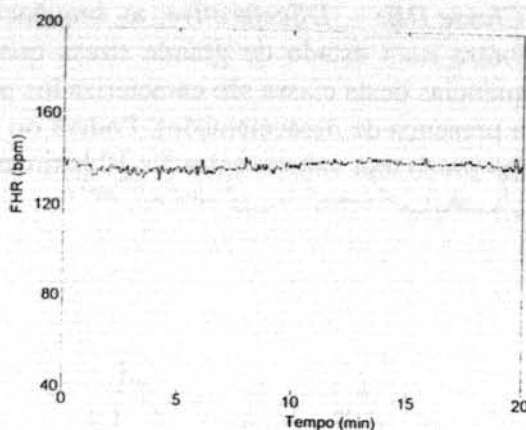


Figura 2-9: Sequência FHR da classe FS

• **Classe S:** *Sinusoidal*, classe idêntica à classe FS, à exceção de as sequências desta classe possuírem uma variabilidade a longo prazo superior, ou seja, 5 ou mais batimentos por minuto. Exemplo na Figura 2-10.

• **Classe SH:** *SHifted*, as sequências FHR desta classe observam-se, por exemplo, nos períodos pós stress, quando existe uma recuperação bastante lenta para a situação normal. Sequências desta classe podem ou não apresentar acelerações e/ou desacelerações. Observa-se a presença de variabilidade a curto prazo e a longo prazo tem valores entre 5 e 10 batimentos por minuto. A característica que mais se evidencia é, no entanto, a presença de intervalos temporais bastante extensos, durante os quais a frequência cardíaca varia de forma praticamente constante. Exemplo na Figura 2-11.

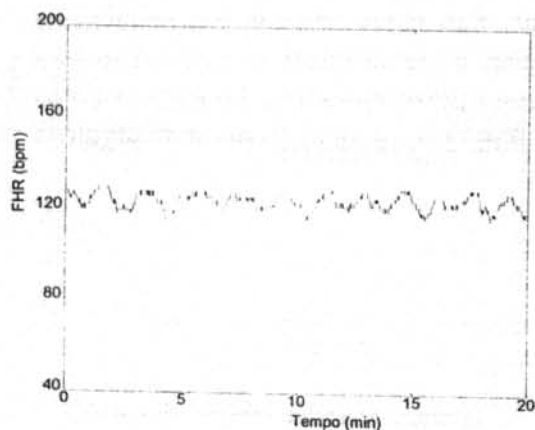


Figura 2-10: Sequência FHR da classe S

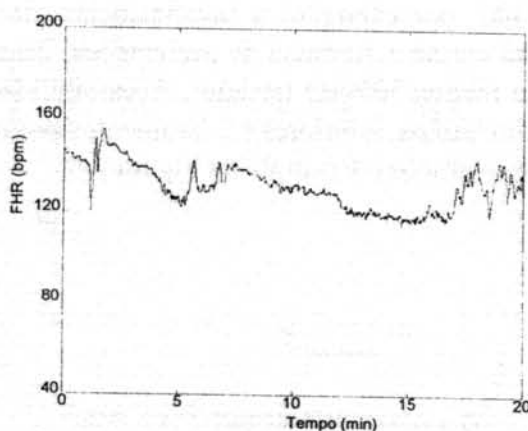


Figura 2-11: Sequência FHR da classe SH

• **Classe TA:** *Tachycardic*, as sequências FHR desta classe indicam que o feto se encontra numa situação fisiológica anormal, por exemplo, um estado febril, de que resulta um estado de taquicardia. As características desta classe são idênticas às da classe B, com exceção da localização da linha base que se deslocou para valores superiores a 150 batimentos por minuto. Exemplo na Figura 2-12.

- **Classe BRA:** *BRAdicardic*, classe de sequências FHR correspondentes a uma situação fisiológica anormal do feto. Nesta situação os fetos apresentam patologias cardíacas, em geral bloqueios aurículo-ventriculares. Sequências desta classe são caracterizados por uma linha base com valor inferior a 110 batimentos por minuto, ausência ou, menos frequentemente, presença de acelerações e desacelerações, ausência de variabilidade a curto prazo e uma variabilidade a longo prazo inferior a 10 batimentos por minuto. Exemplo na Figura 2-13.

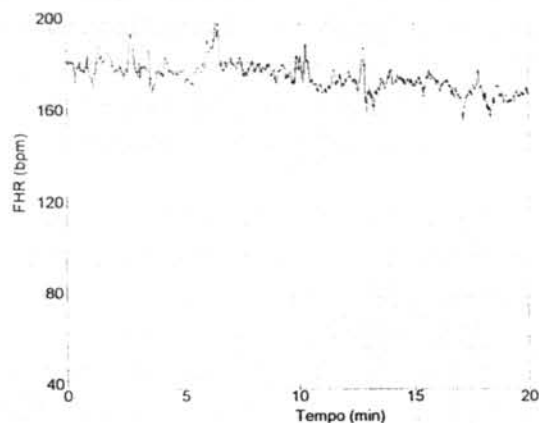


Figura 2-12: Sequência FHR da classe TA

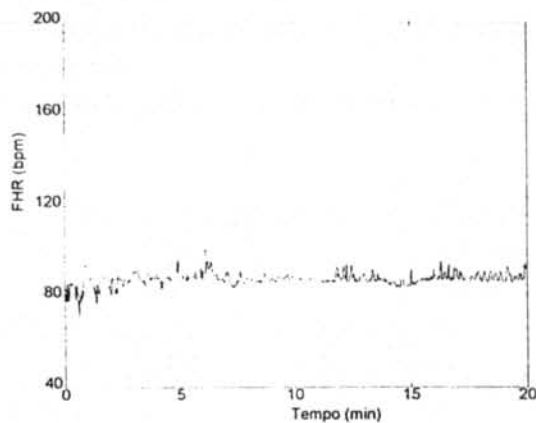


Figura 2-13: Sequência FHR da classe BRA

Algumas das formas da sequência FHR observadas, em especial as pertencentes à classe BRA, são interpretadas à luz do modelo para o coração fetal representado na Figura 2-14. Este modelo considera dois centros de estimulação dos músculos cardíacos: o principal, localizado nas aurículas, induz uma frequência de batimento normal entre 110 e 150 batimentos por minuto e o outro, localizado nos ventrículos, induz batimentos a uma frequência aproximadamente metade da anterior. O bloqueio aurículo-ventricular é interpretado como uma falha do centro principal de estimulação, mantendo-se apenas estimulação dos ventrículos pelo segundo centro, resultando uma sequência FHR com frequência de batimento inferior ao valor normal.

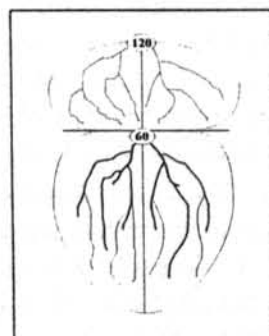


Figura 2-14: Centros de estimulação do coração fetal

Nas figuras anteriores apresentaram-se sequências FHR classificadas como casos puros. No entanto, é habitual que no período de 20 minutos a sequência apresente características de diversas classes, em especial das classes A, B e C, com o feto a atravessar as diversas fases do sono. Esta observação demonstra o carácter não estacionário destas sequências como sinal biológico que é. Esta característica dificulta a análise das sequências, pelo que no presente estudo se utilizaram apenas aquelas que, por inspecção visual, se verificou pertencerem a uma única classe.

2.3 Método de reconstrução de sequências FHR

A função de autocorrelação aplicada a sinais de Doppler, proposta por Tuck 1982 para estimação do período de batimento, calcula, nalguns casos, períodos errados, cujo valor é o dobro ou metade do respectivo valor real. Tuck propôs um procedimento para detecção e eventual correcção destes períodos, observados na sequência FHR como artefactos.

No entanto, muitos monitores fetais que utilizam esta técnica para estimação do período de batimento não fazem qualquer validação do respectivo período. Desta forma a sequência FHR que fornecem apresenta, em geral, um elevado número de artefactos, como ilustrado na Figura 2-15.

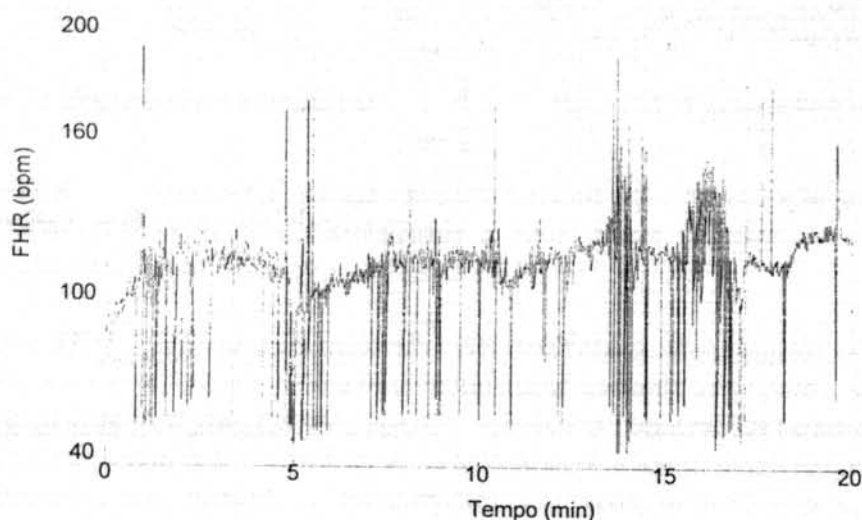


Figura 2-15: Sequência FHR com artefactos

Porque o sinal de Doppler original não está disponível, a correcção da sequência não pode ser realizada utilizando as ideias propostas por Tuck, devendo ser realizada exclusivamente com base na informação disponível na sequência FHR. Já foram propostos diversos métodos para realizar esta correcção (Dawes 1990, Marques de Sá 1992).

A detecção de artefactos é semelhante em ambos os métodos: detecção de descontinuidades da sequência FHR superiores a determinado limiar, e classificação das amostras de uma janela, centrada nas amostras da descontinuidade, como não válidas. No entanto, a correcção é substancialmente diferente.

Como se refere na página 6, Dawes utiliza uma sequência auxiliar. Para efectuar a correcção de artefactos a construção dessa sequência é executada em duas fases. Na primeira, detecta as amostras da sequência auxiliar cujas épocas associadas contêm pelo menos uma amostra válida e determina o respectivo valor pela média das amostras válidas da época. Na segunda fase, estima o valor das épocas que contêm apenas amostras não válidas por interpolação dos valores em épocas adjacentes calculadas na primeira fase. Desta forma garante-se que a duração temporal da sequência é mantida constante.

Na correcção proposta por Marques de Sá, as amostras não válidas não são utilizados nos procedimentos de análise da sequência FHR, podendo considerar-se que foram removidas da sequência. Esta abordagem traz algumas limitações à análise da sequência, em especial à utilização de análise espectral, pois altera-se a duração do sinal e consequentemente a frequência média de amostragem. A substituição dessas amostras por amostras interpoladas de

amostras válidas adjacentes também não é correcta, pois neste caso a duração temporal da sequência corrigida continua a não ser idêntica à duração original.

Com o objectivo de melhorar este último método de correcção, desenvolveu-se um novo método que se passa a descrever.

Nos parágrafos seguintes utilizam-se as seguintes definições:

- Sequência FHR real: é a sequência que se pretende determinar.
- Sequência FHR adquirida: é a sequência medida pelo monitor fetal.
- Sequência FHR estimada: é a sequência obtida da sequência adquirida por correcção dos valores das amostras pertencentes aos artefactos.
- Sequência FHR corrigida: é a sequência obtida pelo método proposto. É uma estimativa da sequência real.

Este método baseia-se na correcção de três tipos básicos de artefactos que se verificou serem as componentes de qualquer artefacto da sequência FHR. Estes artefactos base, já apresentados em Felgueiras 1996, são denominados *half counting*, *double counting* e *spikes*.

Adoptando a representação gráfica utilizada na Figura 2-1 para uma sequência FHR constante, os esquemas respeitantes aos artefactos base citados, ocorrendo numa sequência FHR real constante, estão representados nas Figura 2-16, Figura 2-17 e Figura 2-18, respectivamente. Além de manter a “notação” adoptada na Figura 2-1, nestas figuras a seta com um comprimento inferior indica um batimento não detectado pelo monitor e a seta a tracejado indica a detecção pelo monitor de um artefacto, ruído no sinal de Doppler, a que se associa, consequentemente, um batimento inexistente. Relativamente à sequência FHR, os quadrados pretos indicam amostras correctas da sequência, como anteriormente, enquanto os quadrados brancos marcam amostras da sequência real que não pertencem à sequência adquirida. Os quadrados cinzentos representam amostras da sequência adquirida que pertencem a um artefacto.

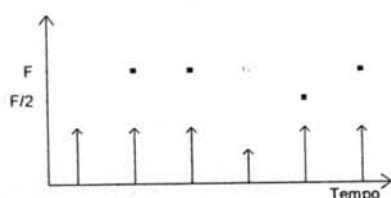


Figura 2-16: Half Counting

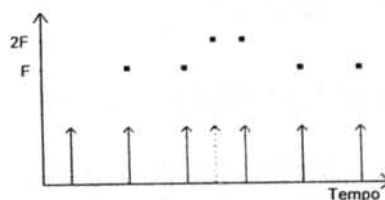


Figura 2-17: Double Counting

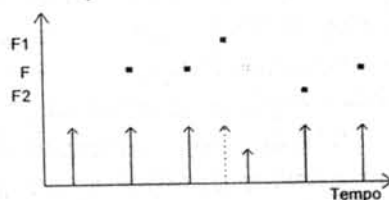


Figura 2-18: Spikes

Assim, um artefacto do tipo *half counting* tem origem quando um batimento não é detectado. Desta forma o período de batimento será o dobro do valor real e, consequentemente, a frequência de batimento terá um valor que é metade do seu valor real.

Note-se, também, que o número de amostras observadas durante a ocorrência do artefacto é metade do seu valor real, por isso este tipo de artefacto se denomina *half counting*.

Um artefacto do tipo *double counting* terá origem quando o monitor fetal detectar um artefacto, no sinal de Doppler, que classifica como um batimento. Assim, o período de batimento será dividido em duas partes, que no caso de serem iguais darão origem a uma frequência de batimento com o dobro do seu valor real. Note-se, que o número de amostras observadas durante a ocorrência do artefacto é o dobro do seu valor real, e daí o nome de *double counting* atribuído a este tipo de artefactos.

Um artefacto do tipo *spikes* ocorre quando o monitor regista um batimento extra, correspondente a um artefacto do sinal de Doppler, e de seguida não detecta um batimento verdadeiro, ou vice-versa. A sequência FHR conterá, desta forma, duas amostras com valores incorrectos: uma com valor superior e outra com valor inferior à frequência de batimento real, ou vice-versa. Neste caso o número de amostras na sequência real e adquirida é igual.

Estas observações são resumidas na conclusão:

Para uma sequência FHR real constante

$$f(i) = F, \quad i = 1, \dots, L,$$

as N amostras da sequência adquirida pertencentes a um artefacto: $f^*(j)$, $j = 1, \dots, N$, satisfazem a relação de continuidade temporal

$$\sum_{j=1}^N \frac{1}{f^*(j)} = \frac{M}{F}, \quad \text{com } M \text{ inteiro.}$$

Pelo que as N amostras pertencentes ao artefacto devem ser substituídas por

$$M = \sum_{j=1}^N \frac{F}{f^*(j)} \quad (1)$$

amostras de valor F , no processo de correcção da sequência FHR.

Desta forma a correcção da sequência FHR deve incluir além da correcção dos valores das amostras pertencentes aos artefactos, a correcção do respectivo número de amostras.

A fórmula deduzida para a determinação do número efectivo de amostras (M) é válida quer para cada um dos artefactos base referidos, quer para qualquer combinação destes, originando artefactos compostos, desde que se verifique a constância da sequência real.

No entanto, a análise de artefactos compostos pode ser simplificada utilizando a seguinte observação acerca da forma dos artefactos base. Amostras de um artefacto composto pertencentes a artefactos do tipo *half counting* contribuem com parcelas inteiras, iguais a 2, para o cálculo de M , enquanto amostras de outro tipo contribuem com parcelas fraccionárias. Assim amostras do tipo *half counting* podem considerar-se artefactos isolados e ser corrigidas sem informação adicional das restantes amostras do artefacto.

A hipótese de constância da sequência real necessita ser relaxada pois como ilustrado em secções anteriores, a sequência FHR não é constante. Assim, o valor de F , da sequência real constante utilizada em (1) para cálculo de M , deve ser substituído pela média harmónica das

respectivas amostras da sequência estimada pois não se dispõe da sequência real. No entanto, esta fórmula de cálculo utiliza $(2n+2)$ multiplicações/divisões e $2n$ adições/subtrações para artefactos com n amostras, pelo que na implementação actual o valor constante F é substituído pela própria sequência estimada, calculando M pela soma da razão das sequências estimada e adquirida, efectuando apenas n multiplicações/divisões e n adições/subtrações. Esta aproximação é válida desde que a amplitude da sequência estimada nas amostras do artefacto seja pequena.

Em resumo as instruções para correcção de uma sequência FHR são:

- ⇒ Determinação da sequência estimada (F_e), pela detecção e correcção dos valores dos artefactos da sequência adquirida (F_a).
- ⇒ Determinação da razão das séries F_e e F_a ($Q = F_e/F_a - 1$).
- ⇒ Detecção dos quocientes Q (Figura 2-19) que diferem de:
 - ◊ 1 menos que δ - identificam as amostras válidas da sequência adquirida,
 - ◊ 2 menos que 2δ - identificam amostras pertencentes a artefactos do tipo half counting,
 - ◊ 3 menos que 3δ - amostras da sequência adquirida geradas quando o monitor fetal não detecta dois batimentos consecutivos, situação que se verificou ocorrer nalgumas sequências,
 e atribuição a essas amostras do número efectivo de amostras M : uma, duas e três, respectivamente.
- ⇒ Detecção de grupos de amostras consecutivas, ainda não consideradas, determinação e atribuição do número efectivo de amostras igual ao arredondamento da soma total dos respectivos quocientes.
- ⇒ Determinação da sequência corrigida por substituição de cada amostra simples pelo número efectivo de amostras com o mesmo valor e substituição de cada grupo de amostras por outro grupo com o número correcto de amostras obtido por interpolação do grupo original.

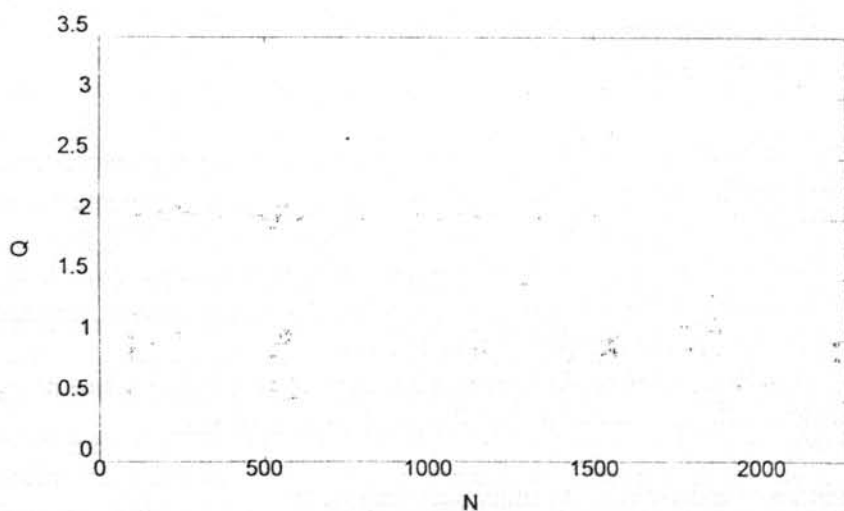


Figura 2-19: Quociente das sequências estimada e adquirida para as amostras dos artefactos da sequência da Figura 2-15

O valor de δ quantifica as diferenças entre a sequência adquirida e estimada, tendo sido fixado em 0.05, a que corresponde um valor absoluto de 6 bpm para o valor de 120 bpm da sequência adquirida. Para uma amostra pertencente a um artefacto do tipo *half counting* com valor 60 bpm o algoritmo proposto anteriormente considera a amostra como um artefacto daquele tipo se o valor da sequência estimada para aquela amostra pertencer ao intervalo [114,126] bpm.

Descrito o procedimento para o cálculo da sequência corrigida a partir das sequência adquirida e estimada, falta descrever o algoritmo utilizado na determinação da última.

A primeira abordagem baseou-se na aplicação de filtros de mediana, pois os artefactos da sequência adquirida manifestam-se essencialmente como ruído do tipo impulsional, com largura de duas amostras no caso de *double counting*, pelo que a aplicação de um filtro de mediana de ordem 5 detecta e corrige artefactos base separados por um mínimo de 3 amostras. No entanto, a presença de artefactos compostos cada vez mais complexos exigem, para um processamento adequado, filtros de mediana de ordem sucessivamente superior. A dificuldade de determinar uma ordem óptima do filtro conduziu ao abandono desta abordagem. Na mesma linha de acção, foi testada a aplicação de filtros morfológicos (abertura e fecho), mas as dificuldades na escolha da dimensão do elemento estruturante revelaram-se semelhantes às surgidas anteriormente.

A abordagem actual é semelhante às referidas no início desta secção, dividindo-se o procedimento de determinação da sequência estimada em duas componentes: detecção de artefactos e correcção de artefactos.

A primeira utiliza a série dos incrementos da sequência adquirida com extensões iniciais e finais, determinadas, respectivamente a partir do histograma dos primeiros e últimos cinquenta valores da sequência adquirida. Estas extensões têm como objectivo gerar condições fronteira adequadas ao processamento e permitir a detecção de artefactos no início e fim de cada sequência. As amostras desta sequência são validadas utilizando a regra de decisão:

$$\left\{ \begin{array}{l} d(i) = |f(i+1) - f(i)| \\ \text{Se } d(i) > D, \text{ as amostras } i \text{ e } i+1 \\ \text{são não válidas} \end{array} \right. \quad (2)$$

O valor do desvio máximo permitido, D , era determinado nas primeiras versões a partir do histograma da sequência $d(i)$ (Felgueiras 1996). Actualmente, considera-se um valor constante, igual a 10 batimentos por minuto, como em Marques de Sá 1992, por se verificar que um valor D dependente das sequências não trazer mais valias, em especial no caso de sequências com baixa variabilidade a curto prazo, mostrando-se, por vezes, inadequado por ser calculado utilizando informação dos próprios artefactos.

As discontinuidades da sequência são codificadas utilizando o início e fim de cada série contínua de amostras não válidas e interpretadas de três formas distintas:

- Descontinuidades “verdadeiras” da sequência adquirida.
- Artefactos da sequência.
- Limites de artefactos.

Exemplos destes três tipos de discontinuidades da sequência particular em estudo estão ilustrados nas Figura 2-20, Figura 2-21 e Figura 2-22. Os símbolos “o” representam o início de discontinuidades, os “x” as amostras finais de cada discontinuidade, enquanto na terceira figura os “+” sobrepostos aos símbolos anteriores indicam os limites de artefactos que foram “fundidos”.



Figura 2-20: Descontinuidades "verdadeiras"

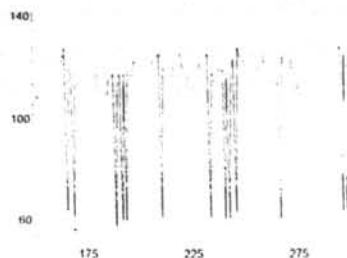


Figura 2-21: Artefactos da sequência



Figura 2-22: Limites de artefactos

Descontinuidades do primeiro tipo ocorrem, por exemplo, nas diminuições abruptas da frequência de batimento na fase final de acelerações em sequências da classe D (Figura 2-5) ou na fase inicial de desacelerações em sequências da classe DE (Figura 2-7). Consideram-se descontinuidades pertencentes a este tipo aquelas que contêm apenas duas amostras e cujo desvio entre as amostras válidas anterior e posterior à descontinuidade é inferior a $\sqrt{3}D$. A correcção destas descontinuidades consiste apenas na interpolação linear das duas amostras utilizando as amostras adjacentes.

Descontinuidades do segundo tipo coincidem com um artefacto da sequência. São classificadas desta forma por o desvio entre as suas extremidades ser inferior a D , verificando-se, portanto, alguma continuidade entre as amostras anteriores e posteriores à descontinuidade. Posteriormente e numa segunda pesquisa, também se consideram pertencentes a este tipo as descontinuidades com desvio, entre extremos, inferior a $2D$. Ambas as descontinuidades são corrigidas por interpolação linear das amostras adjacentes.

O terceiro tipo de descontinuidades ocorre na presença de artefactos compostos, com patamares “estáveis” de três ou mais amostras, como por exemplo na ocorrência consecutiva de dois ou mais artefactos base do tipo *double counting* ou de três ou mais artefactos do tipo *half counting*. Desta forma considera-se que são válidas algumas amostras pertencentes a artefactos. Este problema poderia ser resolvido utilizando janelas de validação das amostras de dimensão superior; no presente caso utilizam-se janelas de três amostras, como pode ser deduzido da regra de decisão (2). Esta solução conduz, no entanto, a problemas semelhantes aos surgidos na escolha da ordem dos filtros de mediana, além de classificar um número “elevado” de amostras válidas, nas abas das descontinuidades, como não válidas.

A solução implementada é semelhante à proposta por Dawes 1990, e consiste em procurar pares de descontinuidades que definam um artefacto. Apesar de Dawes referir que estes artefactos podem ter duração até 10 minutos, nos casos estudados não foram detectados artefactos com tal duração, pelo que a separação máxima entre descontinuidades que delimitam um artefacto foi fixada em 30 amostras. Considerar um número fixo de amostras neste tipo de processamento não é muito correcto, pois pode considerar-se que um intervalo de 12 segundos com frequência de batimento igual a 190 bpm (38 amostras) não define um artefacto, mas se a frequência de batimento for 60 bpm (12 amostras) será considerado como tal.

A correcção deste tipo de descontinuidades consiste, portanto, na detecção de pares de descontinuidades consecutivas que distam menos de 30 amostras e com desvio entre o início da primeira descontinuidade e o fim da segunda inferior a D . Realiza-se ainda uma segunda passagem, procurando os pares cujos desvios são inferiores a $2D$. Para cada par, as descontinuidade são fundidas e corrigidas por interpolação linear das amostras adjacentes.

Após a correcção, pela ordem indicada, das descontinuidades que satisfazem as condições referidas anteriormente, subsistem, nalguns casos, descontinuidades não corrigidas. Estas identificam-se, essencialmente, com as descontinuidades que não verificam os limites impostos para o desvio máximo entre os extremos dos artefactos e podem ser, ou não, características da sequência real. A sua correcção é, no entanto, efectuada da mesma forma: aplicação do filtro de ordem

$$\begin{cases} f_j^0(1) = \text{sort}(f(j+1)), & 1 = \{-4, -3, \dots, 3, 4\} \\ f^*(j) = \frac{1}{3} \sum_{i=-1}^1 f_j^0(i) \end{cases}$$

A resposta deste filtro é óptima quando o artefacto a corrigir tiver, em janelas de nove amostras, no máximo três amostras com valor inferior ao valor correcto e/ou no máximo três amostras com valor superior, pois calcula-se o valor filtrado ($f^*(j)$) a partir da sequência de nove amostras centradas em j , mas tomando apenas a mediana e os dois valores adjacentes.

Os resultados obtidos com o método descrito, são apresentados na Figura 2-23, para uma sequência particular. O traçado superior representa a sequência adquirida pelo monitor fetal, contendo um elevado número de artefactos do tipo *half counting*. O traçado central ilustra a sequência estimada, observando-se uma boa correcção dos valores das amostras pertencentes aos artefactos. Note-se que a escala das abcissas é temporal, e que esta sequência apresenta uma duração significativamente inferior à duração da sequência adquirida, consequência do elevado número de artefactos do tipo *half counting*. Utilizando o método proposto para determinação da sequência corrigida obteve-se o resultado ilustrado no traçado inferior, onde são evidentes as melhorias relativamente à duração temporal da sequência, obtidas com a correcção do número de amostras. A título de exemplo, refira-se que a sequência original continha 2273 amostras e que a sequência corrigida contém 2357 amostras.

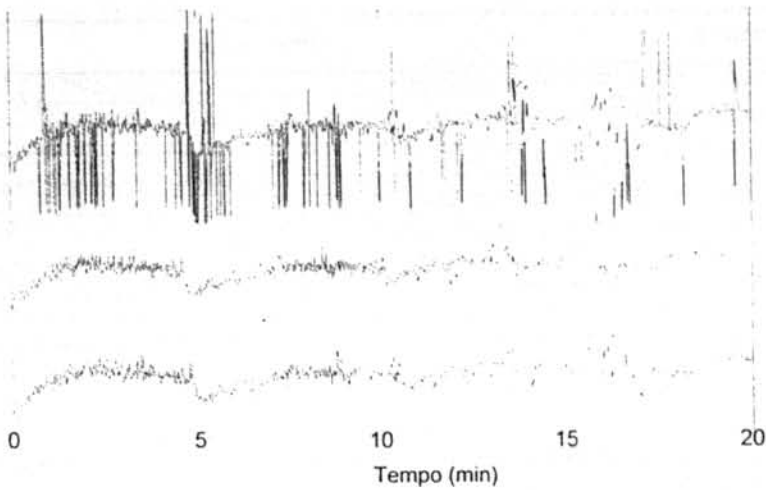


Figura 2-23: Resultados da reconstrução. De cima para baixo: sequência adquirida, sequência estimada e sequência corrigida. Escala das ordenadas arbitrária.

Além desta sequência em particular, foram processados mais 331 sequências. Os resultados obtidos foram bastante satisfatórios. Na Figura 2-24 ilustra-se a diferença da duração temporal das sequências estimada ($T_{ori}-T_{est}$) e corrigida ($T_{ori}-T_{cor}$) relativamente à duração da sequência original.

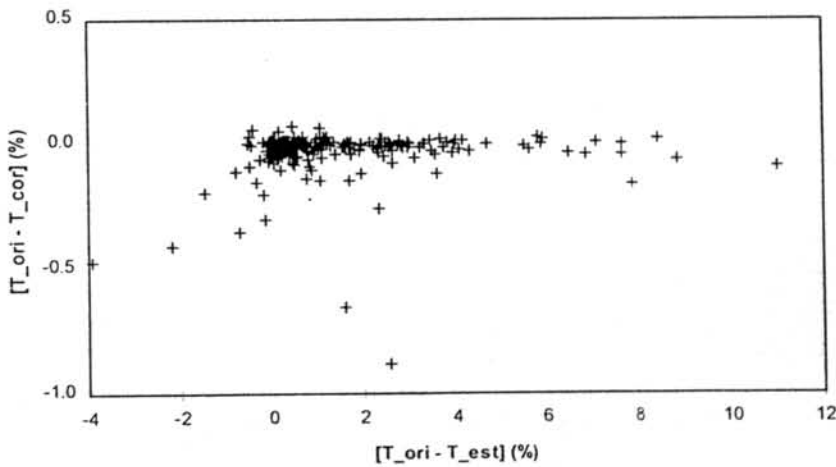


Figura 2-24: Comparação dos tempos das sequências corrigida e estimada.

A distribuição dos pontos, além de ilustrar a significativa melhoria introduzida pela correcção dos número de amostras, permite concluir que a incidência de artefactos do tipo *half counting* - valores positivos de ($T_{ori}-T_{est}$) - é bastante superior à do tipo *double counting*, enquanto o algoritmo de correcção produz sequências com duração, um pouco mas quase sempre, superior à original - valores negativos de ($T_{ori}-T_{cor}$). Estes resultados permitem concluir que a condição de continuidade temporal é verificada no processo de correcção, mas não permite concluir sobre a qualidade da correcção obtida. A verificação dessa qualidade foi obtida por inspecção visual das sequências adquirida e corrigida, sobrepostas na mesma escala, e os resultados são apresentados na Tabela 2-1.

Tabela 2-1: Distribuição dos resultados da correcção de sequências FHR

Resultado	Incidência	Observações
Original	6 %	Sequências sem descontinuidades superiores a D .
Suavizada	12 %	Sequências com descontinuidades, mas que não alteram o respectivo número de amostras (ausência de artefactos do tipo <i>half counting</i> ou <i>double counting</i>). Inclui algumas sequências com ocorrência de artefactos do tipo <i>spikes</i> .
Corrigida	65 %	Sequências com artefactos cuja correcção foi efectuada satisfatoriamente
Incorrecta	17 %	Sequências com artefactos cuja correcção foi efectuada de forma insatisfatória

As sequências cuja correcção foi efectuada de forma incorrecta, incluem, maioritariamente, sequências que têm associados a regiões de elevada variabilidade da sequência real artefactos bastante complexos, de que é exemplo a sequência da Figura 2-25. Esta sequência, classificada por inspecção visual como pertencente à classe DE (Figura 2-7), exhibe, associado às eventuais desacelerações da sequência real, artefactos que impossibilitam a detecção e identificação daquelas.

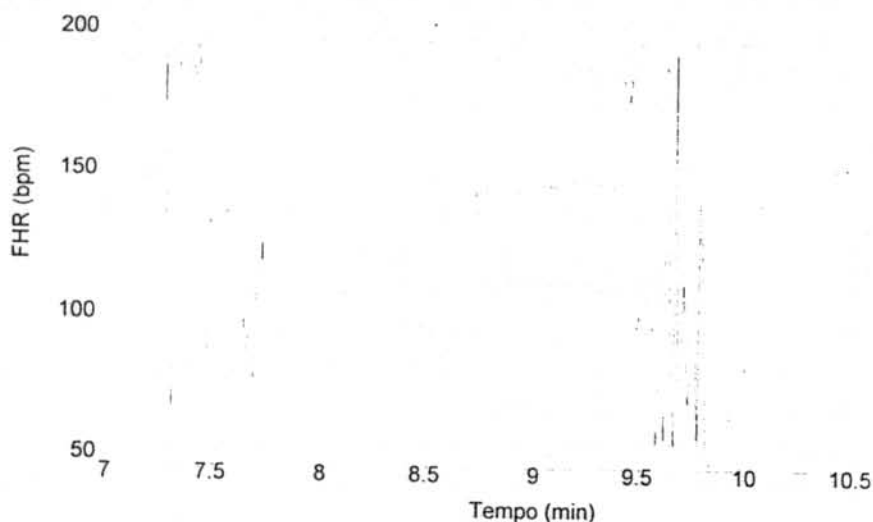


Figura 2-25: Sequência FHR com correcção errada

Pode observar-se uma desaceleração aos 8 minutos, e apesar de desacelerações em sequências deste tipo poderem atingir frequências de batimento bastante baixas (≈ 60 bpm), não é fácil decidir se as amostras com menor frequência de batimento pertencem a uma desaceleração real ou a um artefacto do tipo *half counting*, em especial nos artefactos aos 7 minutos e meio.

Uma outra fonte de resultados incorrectos obtidos com este algoritmo de correcção de sequências FHR, é a união de descontinuidades no sentido de gerar um artefacto de dimensão superior e não a correcção dos artefactos, que ficará impossibilitada se o intervalo entre as descontinuidades que restam for superior aos limite estabelecido de 30 amostras. Este erro resulta do facto de se utilizar uma regra de decisão baseada em valores relativos (diferença entre extremos).

3 Fractais

Este capítulo é dedicado aos objectos fractais, às suas propriedades e caracterização.

Na secção 3.1 ilustra-se o conceito de fractal, como nova classe de objectos matemáticos que modelizam fenómenos naturais. Na secção 3.2 descrevem-se os modelos fractais utilizados no presente trabalho e propõe-se um esquema classificativo para os mesmos. Na secção 3.3 descrevem-se algoritmos de geração de formas fractais, dando especial atenção a fractais temporais. Na secção 3.3.2.2 introduzem-se medidas fractais que permitem a classificação das diversas formas geradas.

3.1 Algumas definições

A modelização de formas naturais por teorias clássicas utiliza objectos geométricos como curvas e superfícies. Os volumes são definidos pelas superfícies que os delimitam, pois considera-se que estes objectos existem num espaço euclidiano tridimensional. Em geral, estas curvas e superfícies não são formas euclidianas, mas quando observadas a escalas suficientemente pequenas parecem-se, respectivamente, com rectas e planos, por isso são adjectivadas de suaves no texto que se segue. Esta propriedade é necessária para usar o cálculo infinitesimal e obter resultados quantitativos, como por exemplo o comprimento de uma curva ou a área de uma superfície.

Porque, nalguns casos, a dimensão euclidiana destes objectos geométricos caracteriza o espaço em que estão embutidos e não os próprios objectos, é conveniente usar uma outra dimensão: a dimensão topológica definida por Poincaré, de forma indutiva, pelas duas proposições (Stewart 1995):

- a) O conjunto vazio tem dimensão topológica -1.
- b) Um conjunto S tem dimensão topológica D_T , se as fronteiras de vizinhanças arbitrárias de qualquer ponto do conjunto tiverem dimensão topológica D_T-1 .

A dimensão topológica assimila coerentemente a noção de dimensionalidade de qualquer conjunto, o que é ilustrado a título de exemplo na Figura 3-1 para conjuntos de dimensão zero, um e dois (ponto, curva e superfície, respectivamente).

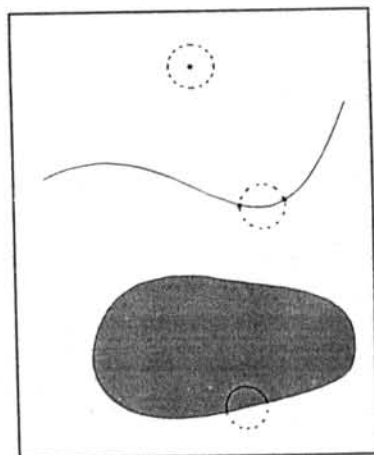


Figura 3-1: Dimensão topológica

Quando se propõe uma teoria é conveniente testá-la, em geral comparando os resultados da teoria com resultados experimentais. Relativamente à modelização de formas naturais, os objectivos dessas experiências podem ser: qual o perímetro da costa portuguesa? Qual a área da árvore pulmonar humana? Verifica-se que os resultados obtidos nestas experiências variam de forma crescente quando se utilizam medidas padrão menores, porque as formas da costa e da árvore pulmonar são rugosas e apresentam estruturas em muitas escalas de ampliação que são detectadas e contabilizadas, para a medida total, apenas por medidas padrão menores que aquelas.

Este resultado, apesar de válido apenas numa gama finita de escalas de ampliação, não é explicado de forma imediata e eficiente pelas teorias clássicas, cujos objectos modelo apresentam um número limitado de comprimentos característicos.

Objectos geométricos mais adequados para modelização destas formas naturais, já eram conhecidos no início deste século, altura em que, por outras razões, alguns matemáticos pesquisaram intensivamente este tipo de objectos. No entanto, seria necessário esperar por 1982 para Mandelbrot apresentar uma nova geometria baseada nesse tipo de objecto

geométrico, por ele denominado *fractal* – da palavra latina *fractus* que significa irregular. Desde então, a geometria fractal tem-se estabelecido como novo paradigma para a descrição de formas geométricas com irregularidades em muitas escalas, em geral e de formas naturais em particular (Mandelbrot 1982).

Apesar dos desenvolvimentos da geometria fractal, a definição de objecto fractal continua a ser uma questão em aberto, pois alguns objectos classificados qualitativamente como fractais não satisfazem os critérios quantitativos das definições propostas. Nos parágrafos seguintes, define-se qualitativamente um objecto fractal, tomando como exemplo a curva de Koch, antes de se propor uma definição baseada em aspectos quantitativos.

A curva de Koch, construída e estudada por Helge von Koch em 1904, está representada na Figura 3-2. A construção desta curva, tal como muitos outros objectos fractais estudados nesta época, é feita por aplicação iterativa de um algoritmo simples.

Como se ilustra na Figura 3-3, o processo de construção da curva de Koch utiliza duas formas geométricas: um segmento de recta ou outra forma poligonal (segmento (a)) designada iniciador e uma forma poligonal (segmento (b)) designada gerador. O iniciador controla a forma geral da curva gerada: por exemplo para se obter uma curva fechada terá, geralmente, de tomar-se uma forma poligonal fechada como iniciador. O gerador controla as características fractais da curva.

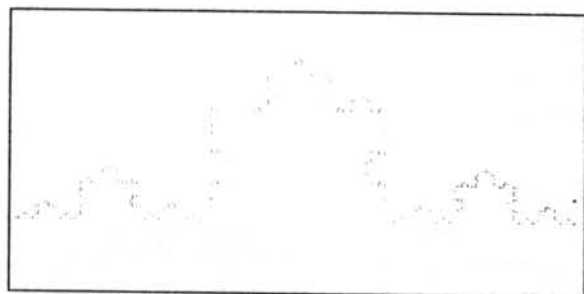


Figura 3-2: Curva de Koch

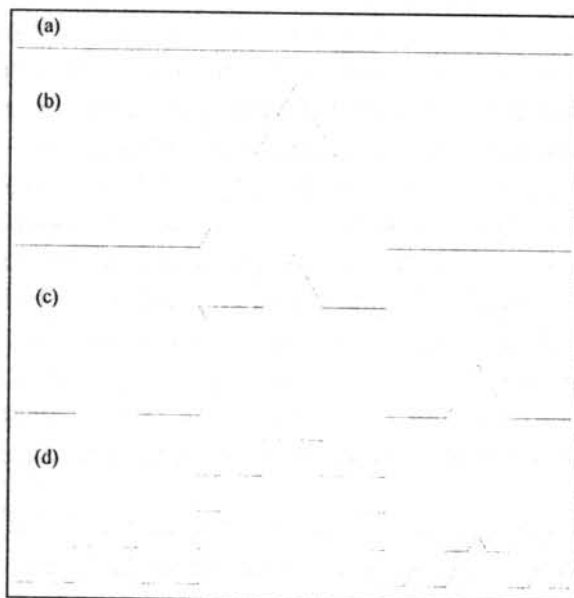


Figura 3-3: Algoritmo de construção da curva de Koch

O algoritmo de construção da curva consiste, simplesmente em substituir cada segmento da aproximação de nível k da curva por uma versão reduzida do gerador, determinando a aproximação de nível $k+1$ de curva. O nível 0 é sempre o iniciador e, no caso presente de um segmento de recta como iniciador, o nível 1 corresponde ao próprio gerador (segmento (b)). Os segmentos (c) e (d) representam os níveis 2 e 3, respectivamente, enquanto a curva da Figura 3-2 ilustra o nível 5 de aproximação, para o qual, a esta resolução, é quase impossível fazer a distinção com o curva de Koch, correspondente ao nível ∞ .

Descrito o processo de construção de objectos fractais, referem-se de seguida algumas das suas propriedades.

O próprio processo de construção é uma característica dos objectos fractais, pois não é possível descrever estes objectos com expressões matemáticas simples como acontece com os

objectos geométricos suaves. Possíveis descrições algébricas de objectos fractais envolvem sempre uma série de Fourier infinita, como para as funções de Weierstrass (Mandelbrot 1982).

Os objectos fractais são não diferenciáveis, ou seja, é impossível definir a tangente ao objecto em qualquer ponto. Aliás, a procura de funções contínuas sem derivada em quase todos os pontos do seu domínio foi uma das ideias que impulsionaram a procura e o estudo deste tipo de objectos.

Fractais como objectos de dimensão topológica D_T têm medidas infinitas associadas a D_T . Por exemplo, considere-se que uma régua de um dado calibre não mede estruturas com comprimento inferior ao seu calibre e, que o iniciador da curva de Koch tem comprimento unitário. O comprimento da curva, quando medido com uma régua calibrada em terços será quatro terços, enquanto que com uma régua calibrada em nonos, valerá 16 nonos, ou seja quatro terços superior ao comprimento medido com a régua “maior”. Continuando a medição com régua de calibre sucessivamente inferior, concluir-se-á que o comprimento da curva de Koch é infinito, apesar de estar contida numa região limitada do plano.

A característica mais importante dos objectos fractais, como a curva de Koch, é a sua auto-semelhança. A definição de auto-semelhança (*self-similarity*), segundo Mandelbrot 1982, faz uso da transformação de semelhança em $\mathbb{R}^E$ – que para um escalar $r > 0$ (razão de semelhança), transforma o ponto $x=(x_1, x_2, \dots, x_E)$ no ponto $T_r(x)=(rx_1, rx_2, \dots, rx_E)$ e um conjunto S no conjunto $T_r(S)$ por aplicação de T_r a todos pontos do conjunto S – e também, da definição de congruência de dois conjuntos, como a propriedade dos dois conjuntos serem idênticos a menos de translações e/ou rotações. A definição é:

Um conjunto limitado, S , diz-se auto-semelhante com respeito ao par (r, N) se S for a união de N conjuntos disjuntos congruentes com $T_r(S)$.

Note-se que a transformação de semelhança é um caso particular de uma transformação de escala (factores de escala iguais a r) ou escalonamento. Pode-se definir auto-semelhança de uma forma menos exacta. Um conjunto diz-se auto-semelhante se existirem subconjuntos que são, após ampliação (escalonamento com $r > 1$), idênticos ao conjunto original.

A curva de Koch é auto-semelhante porque se verifica que cada nível de construção da curva ($b, c, d, \dots$) da Figura 3-3 é a união de $N=4$ réplicas do nível anterior ($a, b, c, \dots$ respectivamente) escalonadas por um factor $r=1/3$ e, portanto, no limite a curva de Koch é auto-semelhante relativamente ao par $(1/3, 4)$. Ou porque, segundo a outra definição, cada quarto da curva da Figura 3-2 resulta após ampliação, por um factor 3, numa curva congruente com a original.

No entanto, os objectos geométricos suaves, podendo ser aproximados a pequenas escalas por segmentos de recta, quadrados ou cubos, também são conjuntos auto-semelhantes. Por exemplo, um segmento de recta de comprimento unitário pode ser considerado a união disjunta de 2 segmentos de recta de comprimento $1/2$, ou de 4 segmentos de comprimento $1/4$, e assim sucessivamente. De igual modo um quadrado de lado unitário é o resultado da união de 4 quadrados de lado $1/2$ ou de 16 quadrados de lado $1/4$, etc. Estes casos e o respeitante a um objecto tridimensional (cubo), são ilustrados na Figura 3-4. Na Tabela 3-1 indicam-se os factores r da transformação de semelhança aplicada ao conjunto original de forma a que este seja a união de N conjuntos transformados, além da dimensão topológica do conjunto.

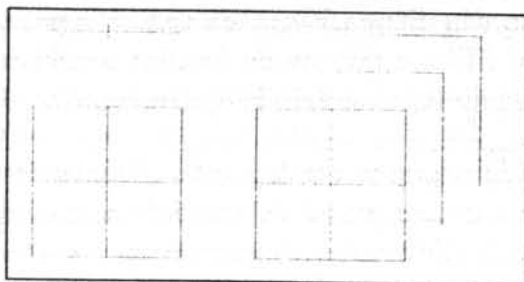


Figura 3-4: Auto-similaridade de formas euclidianas

Tabela 3-1: Pares (r, N) para formas euclidianas

Forma Euclidiana	N	r	D_T
Segmento de Recta	Qualquer	$1/N^{1/1}$	1
Quadrado	Qualquer	$1/N^{1/2}$	2
Cubo	Qualquer	$1/N^{1/3}$	3

Verifica-se que o expoente de N , na determinação de r , está relacionado com a dimensão topológica de cada conjunto. É, portanto, possível definir uma outra dimensão, designada habitualmente por dimensão de semelhança (*similarity dimension*), a partir das seguintes relações:

$$Nr^{D_s} = 1 \quad \text{ou} \quad D_s = \frac{\log(N)}{\log(1/r)}. \quad (1)$$

Apesar de se referir que N pode ser qualquer valor inteiro, isso não é totalmente verdadeiro para os conjuntos limitados que se utilizam. No entanto, note-se que existe alguma liberdade na sua escolha e que essa aumenta quando se considera uma definição mais ampla de conjuntos auto-selhantes.

Um conjunto limitado, S , diz-se auto-selhante com respeito às razões $(r_1, r_2, \dots, r_N)$ se S for a união de N conjuntos disjuntos, cada um congruente com o respectivo conjunto transformado $T_n(S)$.

Neste caso, a dimensão de semelhança de um conjunto auto-selhante com respeito às razões $(r_1, r_2, \dots, r_N)$ é a solução da equação:

$$\sum_{i=1}^N r_i^{D_s} = 1.$$

Na Figura 3-5 representa-se o gerador e o segundo nível de construção de uma curva fractal que é auto-selhante com múltiplas razões de semelhança. Considerando que o comprimento da base é 1, o gerador é formado por 6 segmentos de comprimento $1/3$ e por 5 segmentos de comprimento $1/(3\sqrt{3})$ e, então este objecto fractal será auto-selhante com respeito às razões $((1/3, 6), (1/(3\sqrt{3}), 5))$, porque cada um destes segmentos vai ser substituído por uma versão convenientemente escalonada do gerador. A sua dimensão de auto-selhante será solução da equação:

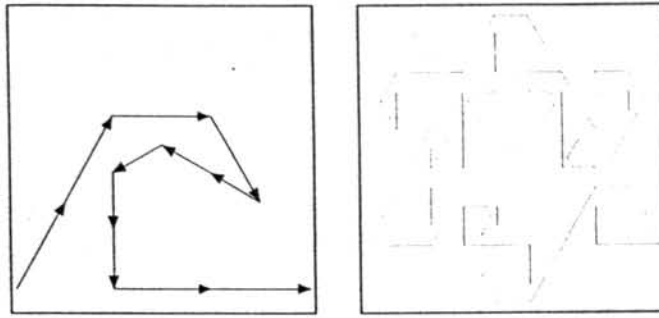


Figura 3-5: Gerador com múltiplos r (esquerda) e segundo nível da curva (direita)

$$6 \times \left(\frac{1}{3}\right)^{D_f} + 5 \times \left(\frac{1}{3\sqrt{3}}\right)^{D_f} = 1$$

Embora a dimensão de semelhança de objectos suaves iguale a dimensão topológica, para objectos fractais, esta tem características diferentes, como se conclui do seguinte exemplo. A dimensão de semelhança da curva de Koch, auto-semelhante em relação ao par $(1/3, 4)$, será por (1) $D_s = \log(4) / \log(3) \approx 1.26$ e da curva da Figura 3-5 é aproximadamente 1.87. O valor desta dimensão tem, mesmo para outros objectos fractais, valores fraccionários superiores à respectiva dimensão topológica. Por isso, é natural que se denomine dimensão fractal, D , e que seja utilizada numa definição de objecto fractal:

Um objecto geométrico é fractal se a sua dimensão fractal, D , for estritamente superior à sua dimensão topológica, D_T .

Além de definir objectos fractais, esta dimensão também permite distinguir fractais com diferentes graus de rugosidade, como se ilustra nas figuras seguintes para aproximações a curvas fractais com 512, 2187 e 1024 segmentos de recta, respectivamente.

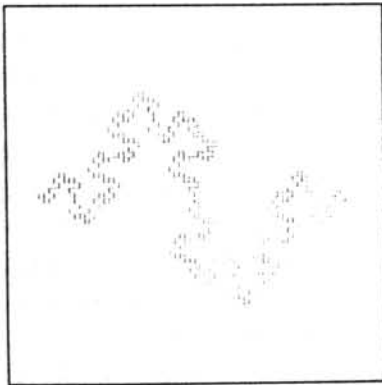


Figura 3-6: $D=1.5$. Nível 3.

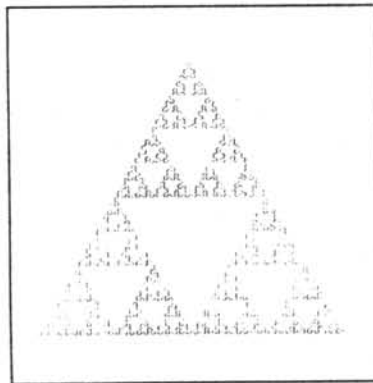


Figura 3-7: $D \approx 1.58$. Nível 7.

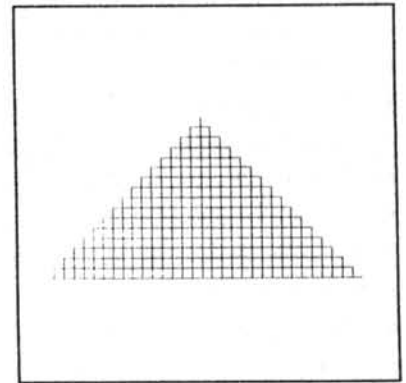
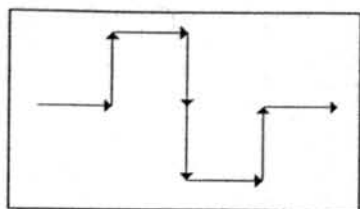
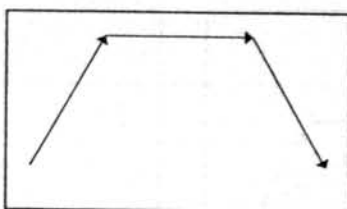
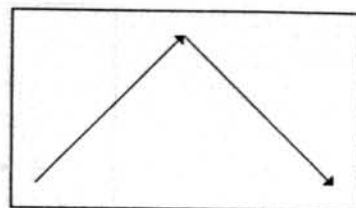


Figura 3-8: $D=2.0$. Nível 10.

Nas figuras seguintes ilustram-se os geradores das curvas ilustradas nas figuras anteriores.

Figura 3-9: Gerador: $N=8$, $r=1/4$.Figura 3-10: Gerador: $N=3$, $r=1/2$.Figura 3-11: Gerador: $N=2$, $r=2^{-1/2}$.

A curva da Figura 3-8, designada curva de Cesaro, pertence à família das curvas de Peano que têm a propriedade de preencherem o plano e é uma excepção à definição de objecto fractal proposta. Esta curva, apesar de ser considerada qualitativamente como fractal – é construída utilizando o mesmo algoritmo, mas com iniciador e gerador adequados –, não o é pela definição de fractal proposta. Como a dimensão topológica caracteriza o conjunto em si, e não o modo como foi gerado, a dimensão topológica da curva de Cesaro e das curvas de Peano é igual 2 tal como a sua dimensão fractal, pelo que estas curvas não satisfazem a condição da definição de objecto fractal.

3.2 Modelos fractais

Apesar de Mandelbrot ter utilizado a palavra fractal para descrever formas geométricas irregulares, os objectos fractais apresentados até ao momento são rugosos mas não podem ser classificados como irregulares. Uma das maneiras para gerar formas irregulares é utilizar o algoritmo apresentado na secção anterior, mas considerando um número elevado de geradores, cuja escolha em cada iteração é feita de forma aleatória. Pode também recorrer-se a modelos do movimento browniano, que se verificou ocorrer em muitos fenómenos naturais. Nesta secção descreve-se e caracteriza-se o movimento browniano como objecto fractal, de seguida descrevem-se os movimentos brownianos fraccionários, introduzidos por Mandelbrot, como extensões do movimento browniano.

3.2.1 Movimento browniano

Enquanto observava partículas de pólen numa suspensão, Robert Brown notou, em 1828, que estas se movimentavam no espaço de forma incessante, irregular e em direcções aleatórias. Além disso, este movimento mantinha as mesmas características quando observado a escalas de tempo distintas, sendo o movimento observado a uma escala de tempo superior a resultante de movimentos idênticos em escalas de tempo inferior. A descrição matemática deste fenómeno só surgiu no início deste século. Mas desde então, o movimento browniano tem sido utilizado com frequência na descrição de muitos fenómenos naturais. Como exemplo, refira-se a sua aplicação em física estatística, para descrição, a nível microscópico, de flutuações e de processos de difusão e a nível macroscópico, em processos de agregação.

A descrição matemática deste movimento utiliza caminhadas aleatórias (*random walks*) e considera o movimento browniano como um limite adequado daquelas caminhadas. Considere-se o caso mais simples: caminhada de uma partícula sobre a recta, com passos de comprimento s dados em instantes múltiplos de T . Além disso, considera-se que os passos são dados de forma independente e com igual probabilidade num sentido ou noutro – seguir para a direita ou esquerda depende, por exemplo, do resultado do lançamento de uma moeda não viciada. A evolução temporal da distância ao ponto de partida toma uma forma semelhante à representada na Figura 3-12.

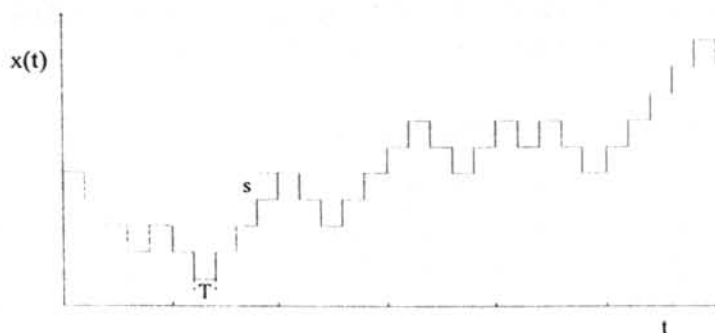


Figura 3-12: Deslocamentos de uma caminhada aleatória

A evolução temporal da posição de uma partícula sujeita a uma caminhada aleatória é uma função aleatória $X(t)$, cujo valor num instante $t=nT$, $x(t=nT)$, é uma variável aleatória. O valor dessa variável é o resultado da soma de n variáveis aleatórias binomiais s_i , com média nula e variância s^2 , pelo que os primeiros momentos desta função aleatória ou processo estocástico, são:

$$E[X(nT)] = 0 \quad , \quad E[X(nT)^2] = ns^2 = \left\lfloor \frac{t}{T} \right\rfloor s^2 .^*$$

Agora, a definição de um limite da caminhada aleatória quando o “período” $T \rightarrow 0$, mantendo t fixo, ou seja $n \rightarrow \infty$, obriga a que $s^2 = \alpha T$, para que a variância do processo não seja infinita. Além disso, de acordo com o Teorema do Limite Central, a função densidade de probabilidade do processo é normal

$$f_X(x, t) = \frac{1}{\sqrt{2\pi\alpha t}} e^{-\frac{x^2}{2\alpha t}} ,$$

ou seja a probabilidade da função $X(t)$ ter valor inferior a x , no instante t , é determinada por uma distribuição normal com média nula e variância αt . Este limite da função $X(t)$ corresponde ao movimento browniano $B(t)$ [†].

Outra abordagem, define o movimento browniano como o integral de um ruído branco gaussiano, ou seja é o limite da soma acumulada de um número infinito de variáveis aleatórias gaussianas e independentes.

* $\lfloor x \rfloor$ representa como é habitual o menor inteiro maior ou igual a x .

† Esta definição pode ser extensível a movimentos em espaços de dimensão superior, desde que os incrementos s_i , variáveis aleatórias multivariadas neste caso, tenham distribuições de probabilidade adequadas.

Assim, o movimento browniano $B(t)$, satisfazendo a condição inicial $B(0)=0$, será um processo estocástico com as seguintes propriedades:

- 1) $E[B(t)] = 0$, ou seja é um processo de média nula;
- 2) $E[B(t)^2] = \nu t$, ou seja a variância do processo é proporcional ao tempo. Dado que a variância não é constante o movimento browniano é um processo não estacionário;
- 3) Atendendo a que o movimento browniano $B(t)$ segue uma distribuição normal que satisfaz as duas propriedades anteriores, então satisfaz, no sentido estatístico, a relação:

$$B(t) = h^{-\frac{1}{2}} B(ht), \forall h > 0. \quad (2)$$

Ou seja as funções densidade de probabilidade dos dois termos da equação anterior são idênticas.

- 4) Para instantes temporais $t_1 < t_2 \leq t_3 < t_4$, os incrementos $B(t_2)-B(t_1)$ e $B(t_4)-B(t_3)$ são independentes;
- 5) Apesar de $B(t)$ ser um processo não estacionário, o processo $\{B(t + \Delta t) - B(t)\}$, com Δt constante é estacionário.

Quando se relaxa a condição $B(0)=0$, as propriedades 1,2 e 3 mantêm-se válidas desde que se refiram aos incrementos do movimento browniano entre quaisquer instantes de tempo e ao respectivo lapso temporal.

A propriedade 3 reflecte a natureza fractal do movimento browniano, pois estabelece uma lei de escalonamento para o movimento browniano. A equação (2) pode ser interpretada de duas formas e permite definir a dimensão fractal dos objectos gerados por movimentos brownianos:

- A primeira resulta na definição de função de Brown que descreve a evolução temporal da posição da partícula, num sistema de eixos com origem na posição inicial da mesma, como na Figura 3-12.
- A segunda corresponde ao movimento browniano, propriamente dito, e constitui o rasto no espaço deixado por uma partícula. Este é difícil de visualizar no caso do movimento browniano sobre a recta que se estudou, pelo que nos parágrafos seguintes consideram-se os movimentos brownianos no plano. A Figura 3-13 ilustra a aproximação com 4096 segmentos de uma realização desse movimento.

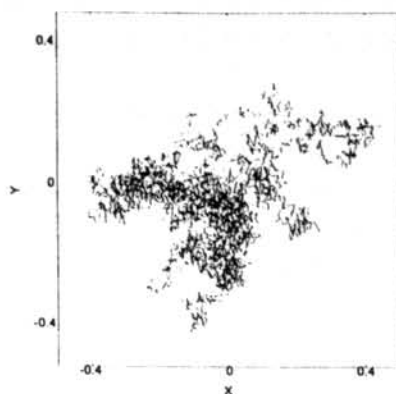


Figura 3-13: Movimento browniano no plano

Considere-se, primeiro o caso do movimento browniano no plano. Neste caso e de acordo com o exposto no início desta secção, o movimento browniano gera formas auto-semelhantes cuja dimensão fractal pode ser determinada utilizando as seguintes ideias.

Como ponto de partida, considere-se a curva de Koch. Utilizando uma variável $t \in [0,1]$ para parametrizar essa curva pelo mapeamento $K(t)$, pode exprimir-se a auto-selhança dessa curva pela relação*

$$K(t) = (1/3)^n K(4^n t), \quad (3)$$

ou seja o primeiro quarto (dezasseis avos, ...) da curva é uma versão escalonada, por um factor $1/3$ ($1/9$, ...), da totalidade da curva. Como a dimensão fractal desta curva é $\log(4)/\log(3)$ e os argumentos dos logaritmos desta expressão estão relacionados com os factores numéricos da equação (3), esta interpretação de relação de auto-selhança permite definir outra variante da dimensão fractal.

Se uma curva mapeada por $C(t)$ satisfizer a relação

$$[C(t+T) - C(T)] = a[C(bt+T) - C(T)]^\dagger,$$

para um conjunto infinito de pares de valores (a, b) , então a curva é um objecto fractal com dimensão fractal

$$D = \lim_{h \rightarrow \infty} \frac{\log(b)}{\log(1/a)}.$$

Com esta definição a dimensão fractal do movimento browniano será

$$D = \lim_{h \rightarrow \infty} \frac{\log(h)}{\log(1/h^{-1/2})} = 2,$$

* Considerar um mapa para este objecto não contradiz uma das propriedades dos objectos fractais, referida anteriormente, pois não é possível determinar a expressão matemática do mapa, apenas se estabelece uma relação que o determina de forma recursiva.

† Esta relação traduz o caso mais geral, como o proposto para o movimento browniano, mas reduz-se às equações (2) e (3) no caso particular de $T=0$ e $C(T)=0$.

e, à excepção do movimento browniano na recta cuja dimensão fractal é 1, todos os movimentos brownianos em espaços de dimensão topológica superior têm dimensão fractal igual a 2.

Além de permitir a definição da dimensão fractal de um maior número de objectos, a interpretação da auto-semelhança proposta permite estabelecer a distinção entre objectos fractais regulares, aqueles que satisfazem a definição de auto-semelhança de forma exacta e têm um espectro discreto de razões de semelhança r , e objectos fractais irregulares, aqueles que apenas a satisfazem de forma estatística e têm um espectro contínuo de razões de semelhança.

As funções de Brown, de que são exemplo os gráficos da Figura 3-14, que representam a evolução temporal das coordenadas do movimento browniano da Figura 3-13, não são curvas auto-semelhantes.

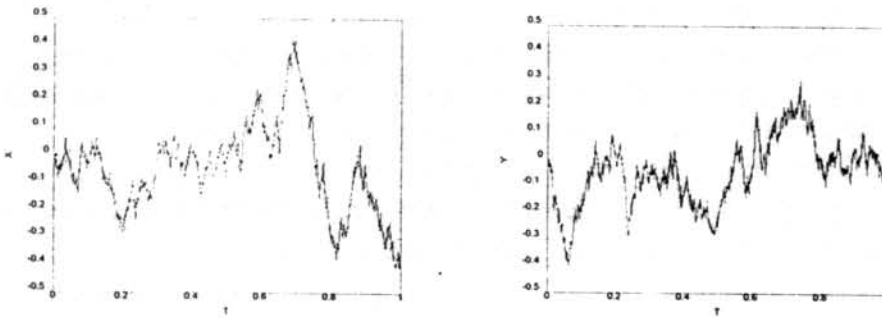


Figura 3-14: Funções de Brown

Porque o seu suporte é um plano com um eixo “temporal” e um eixo “espacial”, a relação de escalonamento indica que os factores de escala a aplicar a cada eixo, para que se verifique a relação de congruência entre os conjuntos original e transformado, são distintos. Este novo tipo de relação de escalonamento é designado por auto-afinidade (*self-affinity*), cuja definição faz uso da definição de transformação afim. Uma transformação em $\mathcal{R}^E$ diz-se afim (*affine*) se transforma o ponto $x=(x_1, x_2, \dots, x_E)$ no ponto $T_r^a(x) = (r_1x_1, r_2x_2, \dots, r_Ex_E)$ com $r = (r_1, r_2, \dots, r_E)$, $r_i > 0$ e pelo menos uma razão r_i distinta das outras. O conjunto transformado $T_r^a(S)$ de S é união de todos os pontos transformados do conjunto S .

Um conjunto limitado, $S \subset \mathcal{R}^T \times \mathcal{R}^E$, diz-se auto-afim com respeito ao par $((r_T, r_E), N)$ se S for a união de N conjuntos disjuntos congruentes com $T_{(r_T, r_E)}^a(S)$.

Os gráficos das funções $X(t)$ e $Y(t)$ são conjuntos contidos em $\mathcal{R} \times \mathcal{R}$ e auto-afins, no sentido estatístico, com razões $(h, h^{-1/2})$.

A dimensão fractal destas curvas não pode ser determinada utilizando, directamente, as definições propostas. No entanto, os conjuntos de zeros das funções $X(t)$ e $Y(t)$ são auto-semelhantes, como ilustrado na Figura 3-15 para os zeros de $X(t)$, pontos inferiores, de $Y(t)$, pontos do meio e zoom na região $t \in [0.12, 0.22]$, pontos superiores, onde é evidente a semelhança entre estes dois últimos conjuntos.

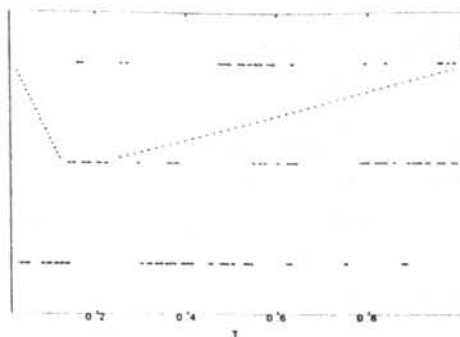


Figura 3-15: Auto-similaridade dos zeros de uma função de Brown

Agora é possível determinar a dimensão fractal do conjunto dos zeros daquelas funções, cujo valor é $1/2$, que indica que este conjunto tem mais estrutura que um ponto, dimensão topológica e fractal igual a 0, mas menos estrutura que uma recta, dimensão topológica e fractal igual a 1.

A dimensão fractal dos gráficos das funções $X(t)$ e $Y(t)$ é $1+1/2=3/2$, pela seguinte regra da geometria euclidiana:

Dois conjuntos S_1 e S_2 de dimensão D_1 e D_2 , respectivamente, num espaço euclidiano de dimensão E , tal que $(E - D_1) + (E - D_2) = E - D$, têm, com probabilidade elevada, uma intersecção não nula com dimensão D se $D > 0$ ou uma intersecção nula se $D \leq 0$.

Esta regra continua a ser válida na geometria fractal, pelo que a dimensão fractal, D , das funções de Brown satisfaz $(2 - D) + (2 - 1) = (2 - 1/2)$ ou seja $D = 1 + 1/2$.

3.2.2 Movimentos brownianos fraccionários (fBm)

Os movimentos brownianos fraccionários, foram propostos e definidos em Mandelbrot em 1968, como uma generalização do movimento browniano, com o objectivo de justificar alguns resultados experimentais obtidos por Hurst.

No meado deste século, Hurst estudou a capacidade de reserva de diversos rios e determinou experimentalmente que as amplitudes desta série temporal durante intervalos temporais, T , eram proporcionais a T^H com $1/2 < H < 1$. Por exemplo, para o rio Nilo esse valor é 0.9. Naquela época os modelos de descargas de rios consideravam que estas eram independentes e distribuídas normalmente – pelo que a capacidade, soma acumulada das descargas, seria modelizada por um movimento browniano – ou eram variáveis dependentes mas descritas uma cadeia de Markov. No entanto, o movimento browniano, estudado na secção anterior, tem uma variância proporcional a T ou, equivalentemente, uma amplitude proporcional a $T^{1/2}$ muito diferente do valor de H determinado, o mesmo se verificando para os outros modelos propostos.

Mandelbrot utilizando a noção de integração e derivação fraccionária deduziu que a lei de Hurst estaria relacionada com a presença de dependência entre descargas com lapso temporais “infinitos”, justificando as diferenças entre os resultados observados e os previstos por teorias clássicas, baseadas na independência ou dependência temporal finita das variáveis.

Procurando novos modelos, Mandelbrot definiu os movimentos brownianos fraccionários (fBm) a partir do movimento browniano $B(t)$:

$$B_H(t_2) - B_H(t_1) = C \left\{ \int_{-\infty}^{t_2} (t_2 - s)^{H-1/2} dB(s) - \int_{-\infty}^{t_1} (t_1 - s)^{H-1/2} dB(s) \right\}, \quad (4)$$

com $0 < H < 1$ e C uma constante, sendo o movimento browniano um caso particular de fBm, ou seja, quando $H=1/2$. O parâmetro H é designado expoente de Hurst.

Os integrais da equação (4), apesar de serem divergentes têm sua diferença finita e definida, podem ser interpretados como uma filtragem. Atendendo à definição do operador de convolução, cada um dos integrais corresponde à convolução (filtragem) de um ruído branco ($dB(s)$) com o *kernel* $s^{H-1/2}$ para $s > 0$ (que define um filtro causal). Além disso aqueles integrais também podem ser interpretados como o integral ou derivada fraccionária de ordem $H-1/2$ de $B(s)$, conforme $H-1/2$ é maior ou menor que 0, ou ao integral fraccionário de ordem $H+1/2$ de $dB(s)$.

O fBm é, tal como movimento browniano, um processo não estacionário cujos incrementos entre os instantes T e $t+T$, com T fixo, são variáveis aleatórias gaussianas com as seguintes propriedades:

- 1) $E[B_H(t+T) - B_H(T)] = 0$;
- 2) $E\{[B_H(t+T) - B_H(T)]^2\} = V_H t^{2H}$;
- 3) $[B_H(t+T) - B_H(T)] = h^{-H}[B_H(ht+T) - B_H(T)] \forall h > 0$, os incrementos do fBm seguem, estatisticamente, uma lei de escalonamento com parâmetro H ;
- 4) Tal com no caso do movimento browniano o processo $\{B(t+T) - B(T)\}$, com t fixo e T variável, é um processo estacionário.

Relativamente à correlação entre os incrementos do fBm em instantes $t_1 < t_2 \leq t_3 < t_4$ distinguem-se três casos, cujas características estão sumariadas na Tabela 3-2, assim como os respectivos efeitos na interpolação e extrapolação de movimentos brownianos fraccionários.

Tabela 3-2: Características dos fBm

$H < 1/2$	$H = 1/2$	$H > 1/2$
Os incrementos em intervalos disjuntos são correlacionados negativamente, ou seja o fBm exibe anti-persistência.	Os incrementos em intervalos disjuntos são independentes.	Os incrementos em intervalos disjuntos são correlacionados positivamente, ou seja o fBm exibe persistência
Derivada fraccionária do movimento browniano ↓↑ Anti-persistência		Integral fraccionário do movimento browniano ↓↑ Persistência
A interpolação do fBm, dados dois valores, é uma função não-linear do tempo.	A interpolação do fBm, dados dois valores, é uma função linear do tempo.	A interpolação do fBm, dados dois valores, é uma função não-linear do tempo.
A extrapolação do fBm a partir de dois valores é no limite temporal infinito igual à média dos dois valores.	A extrapolação do fBm a partir de dois valores é no limite temporal infinito igual ao valor do instante mais próximo	A extrapolação a partir de dois valores tende no limite temporal infinito para um valor infinito
Figura 3-16, Figura 3-19	Figura 3-17, Figura 3-20	Figura 3-18, Figura 3-21

* Estas três propriedades generalizam as propostas para os movimentos brownianos relaxando a condição $B(0)=0$.

Na figuras seguintes representam-se aproximações ao movimentos brownianos fraccionários com 4096 segmentos, assim como exemplos das respectivas funções de Brown fraccionárias (fBf): evolução temporal da abcissa do fBm.

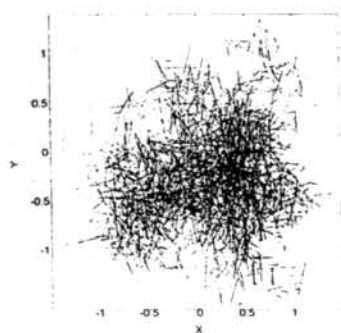


Figura 3-16: fBm $H=0.2$ $D=2.0$

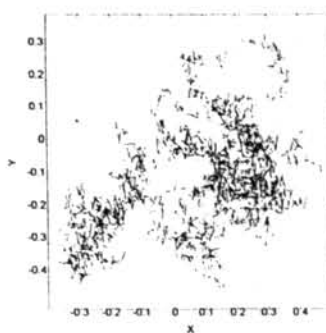


Figura 3-17: fBm $H=0.5$ $D=2.0$

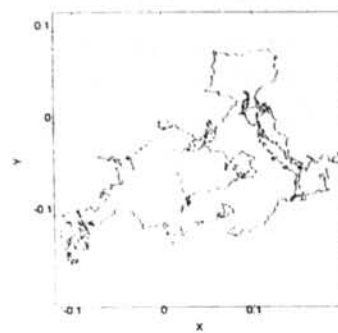


Figura 3-18: fBm $H=0.8$ $D=1.25$

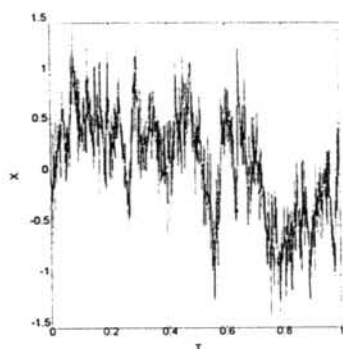


Figura 3-19: Função de Brown
 $H=0.2$ $D=1.8$

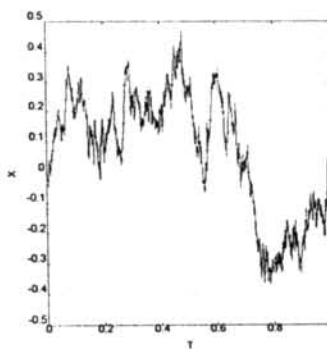


Figura 3-20: Função de Brown
 $H=0.5$ $D=1.5$

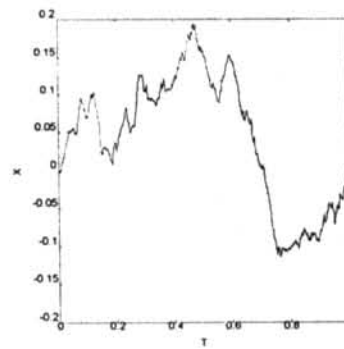


Figura 3-21: Função de Brown
 $H=0.8$ $D=1.2$

Dada a lei de escalonamento da propriedade 3, as dimensões fractais do fBm e da fBf, no plano, são:

fBm	$D = \min(H^{-1}, 2)$
fBf	$D = 2 - H$

A determinação da dimensão fractal dos fBm segue as ideias propostas no caso do movimento browniano, excepto no caso de $H < 1/2$ para o qual a curva se intercepta infinitas vezes e a sua dimensão fractal está limitada superiormente pela dimensão euclidiana do espaço em que está embutida (2, neste caso).

As fBf são, tal como a função de Brown, estudada na secção anterior, objectos auto-afins, cuja dimensão fractal é calculada a partir da dimensão fractal do conjunto de zeros dessas funções. Outra forma de definir o valor $2-H$ para esta dimensão será apresentada em secção posterior.

A densidade espectral das funções de Brown fraccionárias é outra característica importante destes sinais. Sendo sinais não-estacionários, a definição de densidade espectral requer uma análise espectral dependente do tempo, por exemplo o espectro de Wigner-Ville

ou a utilização de transformadas *wavelet*. Estes casos foram estudados por Flandrin 1989, definindo o espectro de uma fBf com expoente de Hurst H :

$$S_{ff}(f) \propto f^{-\beta}, \beta = 2H + 1.$$

3.2.3 Taxinomia

Proposta a definição de objecto fractal e atendendo a que na discussão anterior foram definidos diversos tipos de fractais com propriedade distintas, é conveniente procurar um esquema de classificação de objectos fractais. A classificação proposta é ilustrada na Figura 3-22.

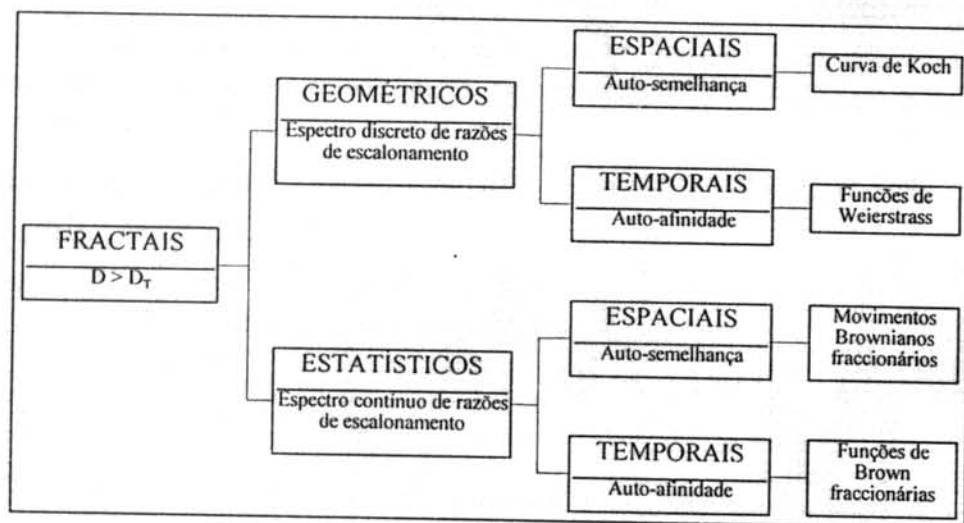


Figura 3-22: Classificação de objectos fractais

Os fractais são divididos em duas classes, estatísticos e geométricos, conforme exibem ou não irregularidades de carácter aleatório, que se manifestam na forma do espectro da razões de escalonamento, como foi discutido anteriormente. Cada uma destas classes é dividida em duas classes, fractais espaciais e fractais temporais, de acordo com o tipo de relação de escalonamento que satisfazem, auto-semelhança ou auto-afinidade.

Os termos espacial e temporal não serão talvez os mais adequados pois, geralmente, estes termos classificam fractais em espaços $\mathcal{R}^E$ e $\mathcal{R}^T \times \mathcal{R}^{E-T}$, respectivamente. Considere-se por exemplo um objecto fractal num espaço tridimensional, isotrópico, classificado como fractal espacial. A sua imagem obtida por uma câmara é, em geral, um fractal temporal pois a cor da imagem tem uma natureza distinta do plano da imagem, que define neste caso um “tempo” com duas coordenadas ($T=2$, $E=1$).

Como modelo para a seqüência FHR consideram-se fractais temporais $\mathcal{R} \times \mathcal{R}$, enquanto os atractores gerados pela representação num espaço de fase daquelas seqüências serão modelizados por fractais espaciais em espaços $\mathcal{R}^E$, com $E \geq 2$.

3.3 Geração de formas fractais

Nesta secção discutem-se os algoritmos utilizados na geração de formas fractais, dando-se especial ênfase à geração de funções de Brown fraccionárias por serem utilizados com modelos para a seqüência FHR.

3.3.1 Geração de fractais geométricos

O algoritmo básico, e utilizado no presente trabalho, para geração de objectos fractais geométricos espaciais já foi apresentado na secção 3.1.

Em termos de implementação calculou-se em cada nível de geração do fractal as coordenadas dos pontos limite de cada segmento. Outras implementações fazem uso de L-sistemas, introduzidos por Lindenmayer com o intuito de modelizar os padrões encontrados em diversos organismos vivos, em especial plantas (Saupe 1988b). Baseada na substituição de símbolos de uma sequência de símbolos, de acordo com determinadas regras, esta implementação tem a vantagem de necessitar uma quantidade consideravelmente menor de memória.

3.3.2 Geração de fractais estatísticos

Nesta secção descrevem-se os algoritmos utilizados na geração de fractais estatísticos, em especial de movimentos brownianos fraccionários (secção 3.3.2.1) e funções de Brown fraccionárias (secção 3.3.2.2).

3.3.2.1 Movimentos brownianos fraccionários (fractais espaciais)

Alguns dos algoritmos apresentados nesta secção são extensões dos algoritmos de geração de fractais geométricos, como o proposto para geração da curva de Koch. Um caso particular destes algoritmos é aquele que utiliza geradores com apenas dois segmentos de recta de comprimentos idênticos, como no caso da curva de Cesaro (Figura 3-11). Estes geradores permitem gerar curvas com qualquer dimensão fractal $1 < D \leq 2$, desde que o valor do comprimento dos segmentos sejam escolhidos de forma adequada. De acordo com a definição de dimensão de semelhança (fractal) o comprimento, r , daqueles segmentos deve ser

$$r = 2^{-D}.$$

Para geradores deste tipo o algoritmo toma cada segmento de recta da aproximação de nível $k-1$, l_{k-1} , e gera os dois segmentos de recta do nível k , l_k , deslocando, perpendicularmente*, o ponto médio do segmento l_{k-1} de um valor Δl_k , determinado pelo valor da dimensão fractal, D , da curva e do nível k .

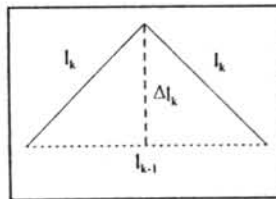


Figura 3-23: Construção do nível k

Para curvas que se iniciam com um segmento de recta de comprimento unitário (iniciador), o valor de Δl_k é determinado pela expressão:

$$|\Delta l_k|^2 = 2^{-2(k+1)/D} - 2^{-2(k/D+1)}. \quad (1)$$

* O sentido do deslocamento obedece a regras bem definidas de forma a gerar as curvas pretendidas

Partindo deste tipo de geradores, os algoritmos para geração de aproximações de movimentos brownianos fraccionários introduzem factores aleatórios de modo a gerar as irregularidades características daquelas curvas.

O primeiro factor aleatório, conforme ilustrado na Figura 3-24, consiste na escolha aleatória e isotrópica da direcção (θ) do deslocamento (a), enquanto o segundo é a escolha aleatória do valor do deslocamento (ΔM) (b), de forma a gerar deslocamentos da forma indicada em (c). Portanto, o algoritmo de geração utiliza em cada passo um gerador distinto e independente dos utilizados anteriormente, determinado por escolha aleatória da direcção do deslocamento e respectiva amplitude.

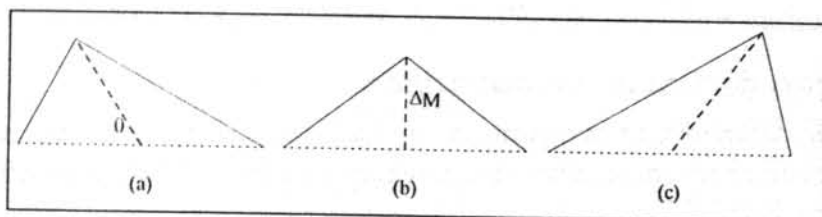


Figura 3-24: Gerador do movimento browniano

A distribuição de probabilidade de ΔM considerada no algoritmo implementado é a proposta por Mandelbrot 1982 na página 245 (“..., the displacement length’s distribution is made Gaussian, and the above formula is made to apply to the mean square displacement.”), ou seja o valor do deslocamento ΔM , na construção do nível k , é o valor absoluto duma variável gaussiana com média nula e desvio padrão idêntico a $|\Delta l_k|$ conforme a equação (1). No entanto, de acordo com o proposto pelo mesmo autor na página 234, referindo-se à geração do movimento browniano (“... the midpoint displacements are made to follow an isotropic Gaussian distribution”, referindo-se ao vector de deslocamento) a distribuição de probabilidade de ΔM utilizada não é a correcta (as curvas geradas exibem um excesso de pequenos deslocamentos).

Este erro não traz, contudo, grande implicações para o presente trabalho a não ser no facto de as funções de Brown das Figura 3-19, Figura 3-20 e Figura 3-21 não o serem de facto, pois não têm as distribuições de probabilidade adequadas, mas os algoritmos de geração destas curvas propostos na secção seguinte produzem resultados mais realistas e correctos. Apesar das curvas geradas pelo algoritmo proposto pertencerem ao plano o caso $H=0.2$ foi gerado considerando $D=1/H=5$ na equação (1), pelo que o movimento gerado (Figura 3-16) se intercepta um elevado de vezes (quase todos os segmentos de recta interceptam outro).

3.3.2.2 Funções de Brown fraccionárias (fractais temporais)

Para geração deste tipo de fractais foram utilizados três algoritmos: método do deslocamento do ponto médio (MPM), método das adições sucessivas (SAM), método da síntese espectral (SSM). Os dois primeiros são adaptações e melhoramentos do algoritmo de geração de movimentos brownianos fraccionários, e utilizam as características estatísticas das funções de Brown. O terceiro utiliza a expressão teórica do espectro das funções de Brown para por transformação inversa de Fourier determinar aproximações daquelas.

3.3.2.2.1 Método do deslocamento do ponto médio

O método do deslocamento do ponto médio (MPM) baseia-se nas propriedades estatísticas das funções de Brown fraccionárias. Considerando uma função de Brown

fraccionária $B(t)$, satisfazendo as propriedades da página 32, a construção de $B(t)$ no intervalo $t \in [0,1]$ processa-se segundo os passos:

- $\Rightarrow B(0)=0$
- $\Rightarrow B(1)=\text{valor}, D_0$, de uma variável gaussiana com média nula e variância V .
- $\Rightarrow$ Dada uma aproximação de $B(t)$ com resolução $\Delta t_{k-1}=2^{-(k-1)}$, $B_{k-1}(n\Delta t_{k-1})$ $n=0, \dots, 2^{k-1}$, a aproximação com resolução $\Delta t_k=2^{-k}$ é construída conforme ilustrado na Figura 3-25,

$$\begin{cases} B_k(2n\Delta t_k) = B_{k-1}(n\Delta t_{k-1}) \\ B_k((2n+1)\Delta t_k) = \frac{B_{k-1}(n\Delta t_{k-1}) + B_{k-1}((n+1)\Delta t_{k-1})}{2} + D_k^n \end{cases}$$

tomando para valor de D_k^n n amostras independentes de uma variável aleatória gaussiana, D_k , com momentos determinados de forma a que média e variância dos incrementos de B_k satisfaçam as propriedades 1) e 2) da página 32.

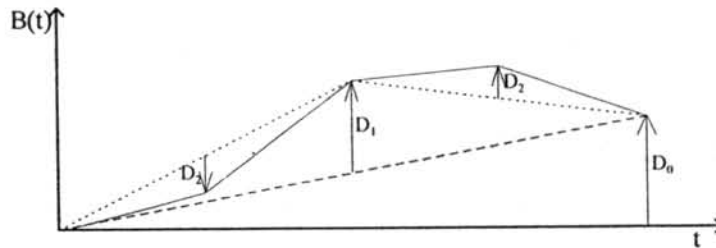


Figura 3-25: O método do deslocamento do ponto médio. Três níveis de aproximação (linha a tracejado, pontead e a cheio).

De acordo com Saupe 1988, a média e variância da variável D_k são:

$$E[D_k] = 0 \quad \text{e} \quad E[D_k^2] = \frac{V}{2^{2kH}} [1 - 2^{2H-2}],$$

sendo H o expoente de Hurst da função de Brown que se pretende gerar.

As aproximações geradas desta forma não são totalmente correctas pois os incrementos da função estimada não pertencem a um processo estacionário, como é próprio das fBf. Este facto resulta de os incrementos, em cada nível de geração, serem independentes, pelo que se verifica, geralmente, que

$$E[(B(t_1 + \Delta t) - B(t_1))^2] \neq E[(B(t_2 + \Delta t) - B(t_2))^2] \quad \text{com} \quad t_1 \neq t_2,$$

sendo, segundo Mandelbrot 1982, o caso mais desfavorável quando $t_1 = n \cdot 2^{-k}$ e $t_2 = (n+0.5) \cdot 2^{-k}$, com $\Delta t = 2^{-k}$, verificando-se que

$$E[(B(t_1 + \Delta t) - B(t_1))^2] = 2^{1-2H} E[(B(t_1 + \Delta t) - B(t_1))^2].$$

O único caso em que não ocorre este facto é quando $H=1/2$, aliás como seria de esperar pela própria definição de função de Brown, com incrementos gaussianos independentes. Nos

outros casos este facto, que constitui uma desvantagem para o presente método, é compensada pela simplicidade do método e a vantagem de ser fácil a obtenção de aproximações com resolução temporal superior a partir de aproximações com resolução inferior.

Nas figuras seguintes ilustram-se aproximações com 1025 pontos de fBf 's, para $H=0.2$, 0.5 e 0.8 , obtidas pelo presente método e utilizando a mesma variável aleatória normalizada, note-se que os valores para 0 e 1 são idênticos.

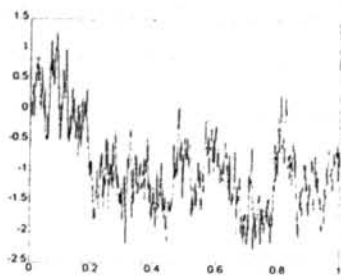


Figura 3-26: MPM, $H=0.2$

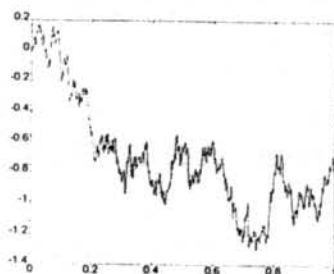


Figura 3-27: MPM, $H=0.5$

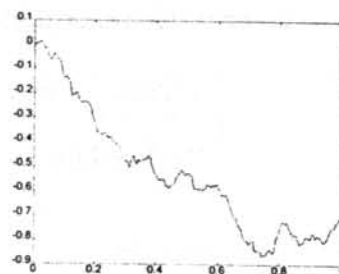


Figura 3-28: MPM, $H=0.8$

Apesar da escala das ordenadas ser distinta são evidentes as diferenças relativamente à rugosidade das curvas, verificando-se qualitativamente que a dimensão fractal das mesmas segue a lei $D=2-H$.

3.3.2.2 Método das adições sucessivas

As desvantagens do método do deslocamento do ponto médio podem ser minimizadas utilizando o método das adições sucessivas (SAM). Este método ao contrário do MPM, que no processo de geração aumenta apenas a resolução temporal da função, aumenta também a resolução dos valores da própria função, ou seja em cada nível de geração além de determinar novos valores, aumentando a resolução temporal, actualiza igualmente os valores já determinados. O método está ilustrado na Figura 3-29, e difere do MPM por deslocar, além do ponto médio, os extremos de cada segmento conforme os passos de execução que seguem.

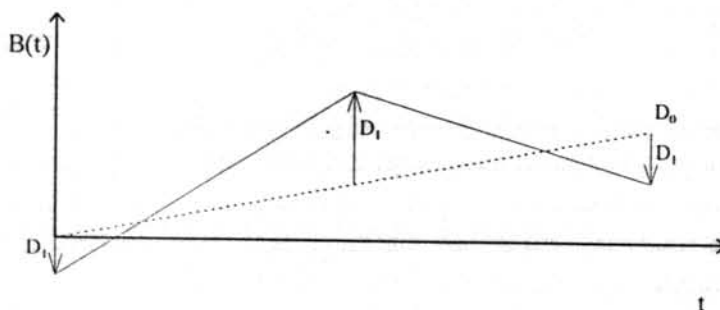


Figura 3-29 O método das adições sucessivas. Dois níveis de aproximação (linha a ponteeado e a cheio).

- ⇒ $B(0)=0$
- ⇒ $B(1)=$ valor, D_0 , de uma variável gaussiana com média nula e variância V .
- ⇒ Dada uma aproximação de $B(t)$ com resolução $\Delta t_{k-1}=2^{-(k-1)}$, $B_{k-1}(n\Delta t_{k-1})$ $n=0, \dots, 2^{k-1}$, a aproximação com resolução $\Delta t_k=2^{-k}$ é construída seguindo as regras,

$$\begin{cases} B_k(2n\Delta t_k) = B_{k-1}(n\Delta t_{k-1}) + D_k^n \\ B_k((2n+1)\Delta t_k) = \frac{B_{k-1}(n\Delta t_{k-1}) + B_{k-1}((n+1)\Delta t_{k-1})}{2} + D_k'^n \end{cases}$$

onde D_k^n e $D_k'^n$ são amostras independentes de uma variável aleatória gaussiana, D_k , com momentos determinados de forma a que média e variância dos incrementos de B_k satisfaçam as propriedades 1) e 2) da página 32.

Tal como no MPM a média da variável D_k é zero sendo a sua variância metade do valor determinado para o método anterior, ou seja

$$E[D_k^2] = \frac{1}{2} \frac{V}{2^{2kH}} [1 - 2^{2H-2}],$$

pois agora a variância dos incrementos da função, em cada nível de resolução é dividida pelos dois incrementos, D_k^n e $D_k'^n$.

Nas figuras seguintes ilustram-se aproximações com 1025 pontos de fBf's, para $H=0.2$, 0.5 e 0.8 , obtidas pelo presente método e utilizando a mesma variável aleatória, mas agora os valores da função no instante $t=1$ são distintos devido à adição de sucessivas amostras. Os valores para $t=0$ foram normalizados a zero por deslocamento vertical da curva.

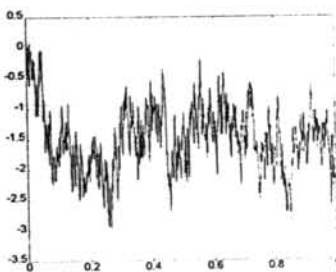


Figura 3-30: SAM, $H=0.2$

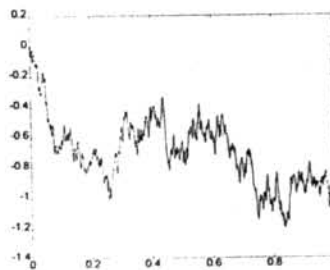


Figura 3-31: SAM, $H=0.5$

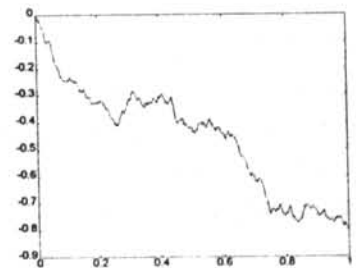


Figura 3-32: SAM, $H=0.8$

3.3.2.2.3 Método da síntese espectral

Outro método de geração de fBf utiliza a densidade espectral destas funções aleatórias para por transformação inversa de Fourier determinar aproximações das mesmas.

Conforme referido na página 34 a densidade espectral desta funções tem a forma geral

$$S_H(f) \propto f^{-\beta}, \beta = 2H + 1.$$

Para gerar uma aproximação da função no intervalo $t \in [0,1[$, com $N=2^k$ pontos amostrados com uma resolução temporal $\Delta t=2^{-k}$, determina-se a transformada discreta de Fourier (DFT)

$$\hat{B}(k\Delta f) = \sum_{n=0}^{N-1} B(n\Delta t) e^{-\frac{2\pi i k n}{N}}, \quad \text{com } \Delta f = 1,$$

de forma a que os coeficientes complexos $\hat{B}(k\Delta f)$ estimem a versão discreta da densidade espectral $S_B(k\Delta f)$. Exprimindo aqueles coeficientes na forma de Euler

$$\hat{B}(k) = a_k e^{-i\theta_k},$$

e atendendo à relação entre a densidade espectral e o quadrado da amplitude da transformada de Fourier então, segundo Saupe 1988, θ_k é uma variável aleatória com valores no intervalo $[0,2\pi[$ e a_k uma variável aleatória que satisfaça

$$E[a_k^2] \propto k^{-\beta}.$$

No entanto, no presente trabalho considera-se, de acordo com os estudos de Higushi 1990, que os valores de a_k são determinados exactamente pela expressão da densidade espectral, mantendo apenas o carácter aleatório da fase θ_k .

Determinados os coeficientes da DFT, a aproximação da função de Brown fraccionária pode ser obtida por transformação inversa de Fourier utilizando uma Fast Fourier Transform (FFT). Neste caso deve ter-se em atenção o tipo de resultado que se pretende, dado que se pretende uma função real os coeficientes $\hat{B}(k)$ devem satisfazer a seguinte relação

$$\hat{B}(k) = \hat{B}^*(N - k),$$

onde X^* designa o complexo conjugado de X . Além disso o factor de proporcionalidade da densidade espectral é determinado utilizando o Teorema de Parseval impondo que a variância dos valores da função estimada seja unitária.

Desta forma o processo de geração de uma aproximação no intervalo $[0,1[$, com $N=2^k$ pontos, para maior eficiência da FFT, segue os seguintes passos:

⇒ Gerar as componentes da DFT:

$$\diamond a_0 = 0$$

$$\diamond \theta_0 = 0$$

$$\diamond \text{ Para } k = 1:N/2-1$$

$$a_k = k^{-(2H+1)/2}$$

$$\theta_k = 2\pi \cdot U(0,1)$$

$$\diamond a_{N/2} = (N/2)^{-(2H+1)/2}$$

$$\diamond \theta_0 = 0$$

$$\diamond \text{ Para } k = N/2+1:N-1$$

$$a_k = a_{N-k}$$

$$\theta_k = -\theta_{N-k}$$

$$\diamond \hat{B}(k) = C a_k e^{-i\theta_k} \text{ com } C = N \left(\sum a_k^2 \right)^{-1/2}$$

⇒ Gerar amostras da função

$$\diamond B(n) = \text{IFFT}(\hat{B}(k)), \quad n = 0:N-1$$

As funções estimadas desta forma são, de acordo com as propriedades da DFT, periódicas com período 1. Este facto constitui uma desvantagem deste método, mas pode ser solucionado facilmente pela estimativa da função num número superior de pontos ($2N$ ou $4N$), da qual se escolhe um segmento com o número pretendido de pontos e se faz o mapeamento para o intervalo $[0,1]$.

Nas figuras seguintes ilustram-se aproximações com 1025 pontos de fBf's, para $H=0.2$, 0.5 e 0.8 , obtidas com este método e utilizando as mesmas amostras da variável uniforme $2\pi \cdot U(0,1)$ para a fase θ_k . Neste caso não se observa a dependência da amplitude da função com o parâmetro H , como nos casos anteriores, devido à normalização da variância das amostras da função.

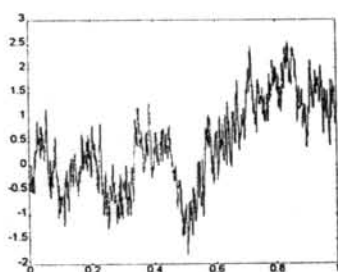


Figura 3-33: SAM, $H=0.2$

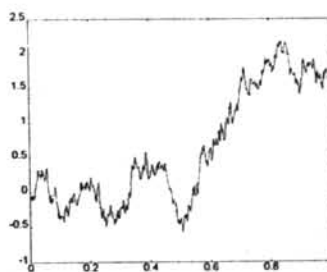


Figura 3-34: SAM, $H=0.2$

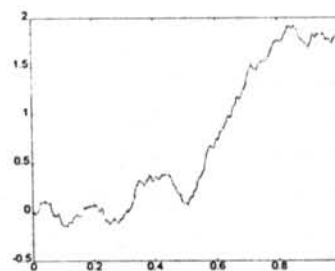


Figura 3-35: SAM, $H=0.2$

Este método sendo mais correcto que os anteriores, tem algumas desvantagens relativamente àqueles. Uma refere-se à complexidade do algoritmo, com fonte no algoritmo da FFT, mas que não se manifesta para o caso dos sinais utilizados (unidimensionais e número de amostras não muito elevado), sendo a outra bem mais restritiva pois impossibilita a utilização das aproximações das funções com resolução inferior na construção de aproximações com resolução superiores como nos métodos anteriores, sendo necessário neste caso a repetição dos cálculos quando se pretende nova aproximação.

Conclui-se aqui a descrição do método de geração de formas fractais utilizadas neste trabalho, discutindo-se na próxima secção os métodos de estimação de dimensões fractais.

3.4 Estimação de dimensões fractais

Nesta secção discutem-se os métodos de estimação da dimensão fractal utilizados neste trabalho. Começa-se por estudar o caso de fractais espaciais, com interesse para as técnicas que serão apresentadas no capítulo Teoria do caos, mas será dada especial ênfase aos métodos de estimação de medidas fractais de fractais temporais e respectivas diferenças com os métodos utilizados para fractais espaciais.

Sem perda de generalidade consideram-se apenas as curvas fractais no plano, portanto objectos fractais cuja dimensão fractal pertence ao intervalo $]1,2]$.

3.4.1 Dimensão box-counting

Uma das propriedades das curvas fractais enunciada na secção 3.1 referia que o comprimento deste tipo de objectos fractais era infinito. Este resultado foi obtido observando a evolução do comprimento da curva medido com régua de comprimento sucessivamente menor. A definição do comprimento $L(r)$ da curva medido com uma régua de comprimento r é, como é usual:

$$L(r) = r \cdot N(r),$$

onde $N(r)$ é o número régua necessárias para definir a curva com nível de resolução r .

No caso fractais geométricos $N(r)$ é determinado exactamente pela dimensão fractal, D , da curva. De acordo com a definição de dimensão de semelhança (fractal)

$$N(r) \propto r^{-D}. \quad (1)$$

Desta forma a relação

$$L(r) \propto r^{-D+1}, \quad (2)$$

que também é válida, em média, no caso de fractais estatísticos, permite estimar a dimensão fractal de uma curva determinando o declive da recta formada pela representação gráfica, em escalas logarítmicas, dos pontos experimentais $(r, L(r))$. Voss 1988 designa a dimensão fractal estimada desta forma por dimensão de compasso (instrumento que permite definir segmentos de recta com um comprimento especificado e qualquer orientação no plano).

Outra forma de estimar a dimensão fractal de uma curva faz uso da equação (1), sendo o valor $N(r)$ determinado pelo número mínimo de circunferência de raio r necessárias para cobrir a curva. Uma outra versão consiste em utilizar quadrados de lado r em vez de circunferências, pelo que é habitual designar a dimensão fractal estimada desta forma por dimensão de caixa (*box dimension*).

No entanto, estes métodos de estimação não são implementados em algoritmos computacionais com facilidade devido a pesquisa que é necessário efectuar percorrendo a própria curva. Um método mais simples, baseado no último método, é ilustrado na Figura 3-36, para o caso de um fractal geométrico (floco de neve de Koch) e um fractal estatístico (fBm com $D=1.5$). Este método estima $N(r)$ determinado o número de caixas de uma grelha com espaçamento r que contêm pelo menos um ponto da curva (na figura grelhas com diferentes espaçamentos são representadas com diferentes tipos de linha, existindo a sobreposição de algumas linhas). A estimativa de $N(r)$ obtida desta forma é geralmente uma estimativa por excesso do valor real, podendo ser determinada mais fielmente tomando a média dos $N(r)$ obtidos para as grelhas com o mesmo espaçamento mas com origens distintas. A dimensão

fractal estimada desta forma é designada, segundo Voss 1988, por dimensão de grelha (*grid dimension*), sendo mais usual a denominação dimensão de contagem de caixas (*box counting dimension*), que será utilizada no resto do texto.

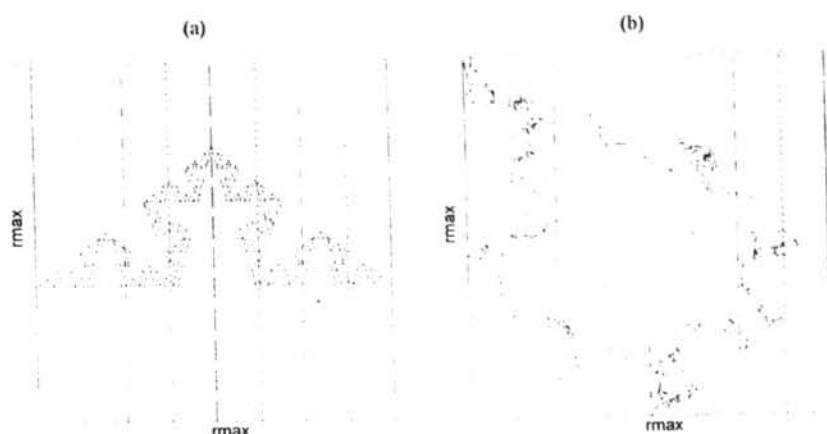


Figura 3-36: Dimensão de grelha ou contagem de caixas de fractais geométrico (a) e estatístico (b) com dimensão fractal $D=1.5$. (r_{max} deve ler-se L_{max})

Apesar de este método ser particularmente interessante quando aplicado a curvas que estão representadas no monitor de um computador aproveitando os pixels da imagem para geração das grelhas, o algoritmo implementado utiliza as coordenadas “verdadeiras” dos pontos da curva. Esta abordagem permite, desta forma, estimar a dimensão fractal de curvas em espaços de dimensão superior a dois, como vai ser requerido pelas técnicas estudadas na secção 4.2.

O algoritmo implementado utiliza árvores binárias de forma a minimizar os requerimentos de memória, pelo que as grelhas utilizadas têm espaçamentos que são sucessivamente divididos por dois, tal como ilustrado na figura anterior. Utiliza apenas uma grelha, não efectuando como exposto anteriormente a média de diversas estimativas de $N(r)$. Esta abordagem pode fornecer estimativas erradas quando o espaçamento da grelha é elevado, mas de forma geral e porque muitos dos objectos que se pretendem medir têm formas aleatórias será pouco provável que para resoluções mais elevadas se considere uma grelha que não quantifique satisfatoriamente a forma do objecto.

Desta forma o algoritmo começa por determinar, utilizando o sistema de eixos euclidianos dos dados, a caixa de largura mínima que contem a totalidade do objecto, na figura anterior designada por r_{max} . Desta forma e sem perda de generalidade pode considerar-se que qualquer ponto $X = (x_1, x_2)$ do objecto pertence ao quadrado unitário $([0,1] \times [0,1])$, pelo que o espaçamento das grelhas utilizadas neste método é determinado por $r = 2^{-k}$.

O próximo passo é fixar o espaçamento mínimo das grelhas $r_{min} = 2^{-k_{max}}$, limitado pela resolução espacial do objecto e transformar os n pontos X_i de acordo com

$$Y_i = \lfloor X_i \cdot 2^{k_{max}} \rfloor,$$

sendo os *overflows* (quando uma das coordenadas de X_i é igual a 1) corrigidos atribuindo-lhe o valor zero. A representação binária (com k_{max} bits) dos inteiros Y_i permite determinar a localização de cada ponto na grelha actual, com espaçamento 2^{-k} , de acordo com a regra

$$\begin{cases} Y_i = (y_{i0}, y_{i1}), & i = 1, \dots, n \\ y_{ij} = b_{ij}^0 b_{ij}^1 \dots b_{ij}^{k_{\max}-1}, & j = 0, 1 \\ g_i^{k+1} = b_{i1}^k b_{i0}^k, & k = 0, \dots, k_{\max} - 1 \end{cases}$$

e defini-la pela sequência de $K+1$ nodos de uma árvore $(0, g_i^1, g_i^2, \dots, g_i^K)$, representando o 0 (zero) o nodo inicial árvore.

Apesar de as árvores quaternárias serem as mais adequadas para o caso em estudo (curva no plano, g_i^K toma valores entre 0 e 3) optou-se por implementar a árvore sob a forma binária conforme exemplificado na Figura 3-37, de forma a que seja imediata a generalização para objectos em espaços de dimensão superior.

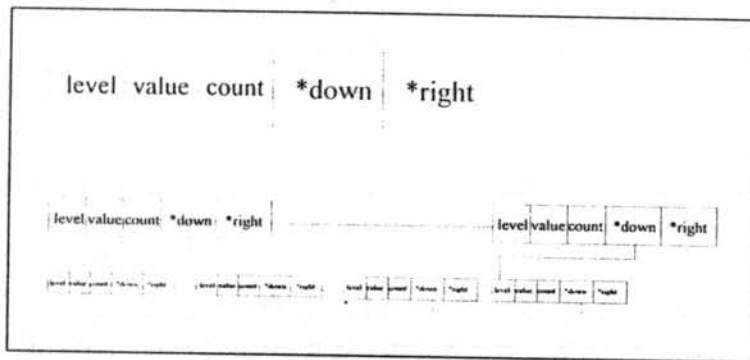


Figura 3-37: Estrutura de dados da árvore

A estrutura de dados dos nodos da árvore contem cinco campos com as seguintes funções.

Tabela 3-3: Campos do nodo da árvore

<i>level</i>	Determina o nível do nodo. Iguala o valor de K da resolução da grelha.
<i>value</i>	Identifica os diferentes nodos em níveis distintos. Iguala o valor de g_i^K .
<i>count</i>	Conta o número de pontos em cada nodo (caixa).
<i>*down</i>	Aponta para nodo no mesmo nível.
<i>*right</i>	Aponta para nodo no nível inferior.

Distribuídos os pontos do objecto pela árvore, de acordo com as regras expostas anteriormente, a determinação do número $(N(r))$ de caixas com comprimento $r = 2^{-k}$ ocupadas por pontos consiste em determinar o número de nodos no nível k da árvore. Representando graficamente o logaritmo de $N(r)$ em função do logaritmo de $1/r$, como ilustrado na Figura 3-38, a dimensão fractal é estimada pelo valor do declive da região linear da curva obtida.

Nesta figura ilustra-se, igualmente, o efeito da resolução das coordenadas dos 1025 pontos que definem os objectos fractais da Figura 3-36. Estão representados os resultados obtidos para codificações com 4, 6, 8 e 10 bits que podem ser identificados pelo ponto em que ocorre a saturação de $N(r)$ (a abscissa iguala o número de bits de codificação utilizado). Nos casos com 4 e 6 bits o valor máximo de $N(r)$ é inferior a 1025 pelo facto de no processo de codificação pontos distintos do objecto serem transformados no mesmo ponto. Os casos com 8 e 10 bits são coincidentes (no caso do fractal geométrico (a)) ou praticamente coincidentes (no caso de fractal estatístico (b)) devido ao número finito de pontos utilizado, desta forma apesar de com 10 bits cada ponto ser definido com maior precisão, o número de

pontos não é suficiente para definir o objecto com precisão superior. No caso do fractal estatístico o número de caixas ocupadas é ligeiramente diferente pois nalgumas regiões do fractal a maior densidade de pontos conduz a uma definição do objecto com maior precisão quando se utilizam mais bits de codificação.

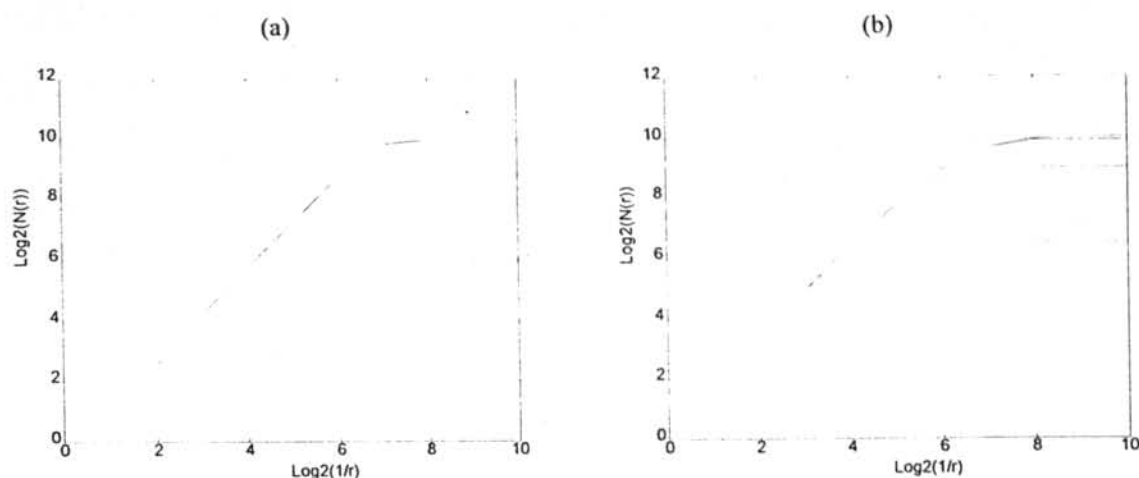


Figura 3-38: Dimensão *box counting* dos fractais da Figura 3-36. As diferentes (4) curvas correspondem a fractais codificados com 4, 6, 8 e 10 bits. A tracejado indica-se segmento com declive correcto ($D=1.5$).

No entanto é evidente a concordância entre o valor estimado da dimensão fractal e o valor real para o fractal geométrico. No caso do fractal estatístico o valor da dimensão fractal deve ser tomado com uma média sobre um elevado número de amostras deste tipo de fractais, no entanto o valor da dimensão fractal poderá ser substancialmente diferente devido erro ao método de geração referido anteriormente.

3.4.2 Dimensões fractais de fractais temporais

O método acabado de descrever apenas é válido quando aplicado a fractais espaciais. No caso de fractais temporais o método da contagem de caixas para determinação da dimensão fractal não é aplicável porque a hipótese de que versões ampliadas de secções de um objecto fractal é semelhante à totalidade do objecto não se verifica neste tipo de fractais. Como exemplo, na Figura 3-39 estão representadas três versões ampliadas de uma função de Brown, para as quais se pode observar que para factores de ampliação superiores a “largura” da linha aumenta. Este facto é quantificado pela variância dos incrementos das funções que como pode observar-se na Tabela 3-4 aumenta com o factor de ampliação.

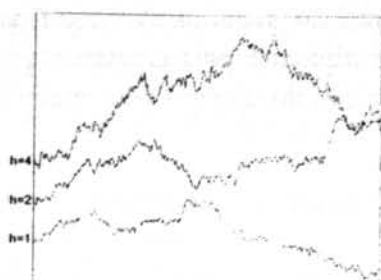


Figura 3-39: Função de Brown ampliada
h=1, 2 e 4 vezes.

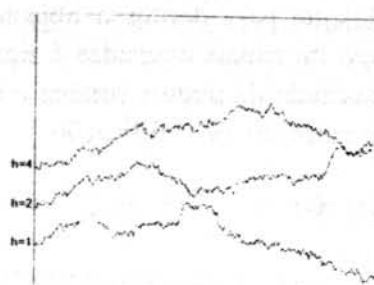


Figura 3-40: Ampliações 1, 2 e 4 vezes de
uma Função de Brown respeitando a
relação de auto-afinidade.

Tabela 3-4: Variância dos incrementos
das funções da Figura 3-39

fBf, h=1	fBf, h=2	fBf, h=4
0.0352	0.0473	0.0600

A relação de escalonamento para fractais temporais é a auto-afinidade. Pelo que é necessário utilizar dois factores de ampliação distintos: $h = 1, 2$ e 4 para o eixo temporal e $h^{1/2}$ para o eixo dos valores da função ($H = 1/2$ para as funções de Brown), como ilustrado na Figura 3-40 verificando-se que agora a “largura” da linha é praticamente idêntica (como confirmado pelos valores da Tabela 3-5).

Tabela 3-5: Variância dos incrementos
das funções da Figura 3-40

fBf, h=1	fBf, h=2*	fBf, h=4*
0.0352	0.0335	0.0300

Apesar da impossibilidade da aplicação prática deste método é, no entanto, possível utilizar este método para definir a relação entre a dimensão fractal, D , e o expoente de Hurst, H , para as funções de Brown fraccionárias (fBf), de acordo com o que vai ser exposto.

De acordo com as propriedades das fBf com expoente de Hurst H , a amplitude da função, ΔV , num intervalo temporal, Δt , satisfaz a relação $\Delta V \propto \Delta t^H$, pelo que a amplitude em intervalos temporais $\Delta t_k = \Delta t / k$ será

$$\Delta V_k \propto \Delta t_k^H \propto \Delta V / k^H,$$

sendo necessários, em média,

$$\Delta V_k / \Delta t_k = \Delta V / \Delta t \cdot k^{1-H}$$

quadrados de comprimento Δt_k para cobrir a função no intervalo temporal Δt_k . O número total de quadrados para cobrir a totalidade da curva será

$$N(k) = k \cdot \Delta V / \Delta t \cdot k^{1-H} = \Delta V / \Delta t \cdot k^{2-H},$$

e a definição de dimensão fractal será, de acordo com a relação $N(r) \propto r^{-D}$,

$$D = 2 - H.$$

Dos muitos métodos para estimar o expoente de Hurst de séries temporais modelizadas por fBf foram implementados três métodos que são descritos na secções seguintes.

3.4.2.1 Expoente espectral

Baseado na forma da função densidade espectral das funções de Brown fraccionárias este método estima o expoente de Hurst determinando o expoente do espectro de potência de uma sequência finita que aproxima a fBf. A estimação do espectro é efectuada por determinação do periodograma de uma sequência $x(n)$ a partir das definições

$$\begin{aligned} I_N(e^{i\omega}) &= \frac{1}{L} |Y(e^{i\omega})|^2 \\ Y(e^{i\omega}) &= \sum_{n=0}^{N-1} y(n) e^{-i\omega n} \quad , \quad y(n) = w(n)x(n) \quad , \\ L &= \sum_{n=0}^{N-1} w^2(n) \end{aligned}$$

ou seja, o periodograma $I_N(e^{i\omega})$ é o quadrado do módulo da transformada discreta de Fourier (DFT), $Y(e^{i\omega})$, de um segmento $y(n)$ obtido da sequência original por aplicação de uma janela $w(n)$ de comprimento N . O factor L é um factor de escala dependente da janela utilizada. Para estimação da DFT utilizou-se uma *Fast Fourier Transform* (FFT) que determina de forma eficiente os valores da DFT para valores da frequência ω da forma

$$\omega_k = \frac{2\pi}{N} k \quad , \quad k = 0:N-1,$$

utilizando-se no texto que se segue a notação $I_N(k)$ e $Y(k)$ para denominar as aproximações ao periodograma e DFT utilizando a FFT.

É possível provar que o periodograma é um estimador não consistente pelo que é usual efectuar a média de diversos periodogramas independentes de forma a diminuir a variância do estimador. O método de Welch implementa estas ideias determinando o periodograma de K segmentos com N amostras obtidas a partir de uma sequência original de M amostras e tomando o periodograma médio

$$I_x^w(k) = \frac{1}{K} \sum_{j=0}^{K-1} I_x^{(j)}(k) \quad ,$$

onde o índice j identifica os segmentos. Em geral o número de segmentos K é determinado pelo quociente M/N , correspondendo a segmentos contíguos, no entanto de acordo com o tipo de janela utilizada para definir os segmentos é possível utilizar segmentos com intersecção não nula mantendo suficientemente válida a condição de independência dos periodogramas e aumentando o seu número que se traduz numa diminuição da variância do periodograma médio. No presente trabalho utilizou-se a janela hanning com sobreposição de 50 % dos segmentos.

Apesar de este método ter sido desenvolvido para estimar a densidade espectral de processos ergódicos, consequentemente estacionários, foi aplicado a sequência que aproximam funções de Brown fraccionárias que são processos não estacionários (neste caso é possível reinterpretar este método como uma estimativa do espectro de Wigner-Ville). Os resultados obtidos para três sequências geradas pelo método da síntese espectral são apresentados na Figura 3-41 onde é possível verificar que o espectro segue uma lei de potência com expoente próximo do valor correcto. No entanto este resultado “qualitativo” não tem correspondência nas unidades do espectro de potência, que será justificado posteriormente, impossibilitando a realização de medidas quantitativas do espectro de potência de fBf's.

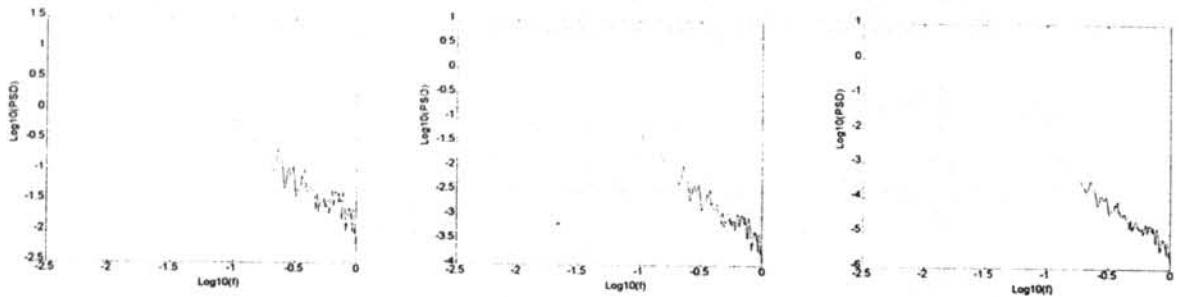


Figura 3-41: Funções densidade espectral de fBf com $H=0.2, 0.5$ e 0.8 estimadas pelo periodograma médio de sequências com 1024 amostras e segmentos com 256 pontos. A frequência indicada está normalizada pela frequência de Nyquist (metade da frequência de amostragem)

Na figura seguinte estão ilustrados a média e respectivo desvio padrão dos expoentes de Hurst determinados, H_d , a partir do expoente espectral, de 64 aproximações de fBf's para cada expoente de Hurst original, H_o , e os três métodos de geração estudados anteriormente. As condições de estimação do espectro de potência são as indicadas na figura anterior. A estimação do expoente espectral é efectuada por determinação do declive da regressão linear do espectro de potência quando representado numa escala bi-logaritmica. Devido à distribuição não uniforme dos pontos da abcissa ($\text{Log}_{10}(f)$) a regressão linear ponderou os pontos com pesos iguais a $1/f$.

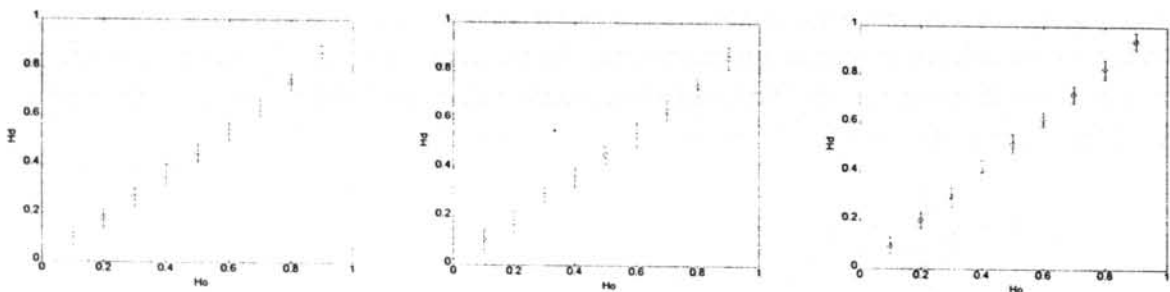


Figura 3-42: Estimação do expoente de Hurst para aproximações de fBf's geradas pelos métodos de deslocamento do ponto médio, das adições sucessivas e da síntese espectral, da esquerda para a direita respectivamente.

Observa-se, em geral, que os expoentes de Hurst determinados são idênticos aos expoentes originais, os maiores desvios não justificados pelo desvio padrão ocorrem para as sequências geradas pelo método do deslocamento do ponto médio podendo dever-se às deficiências do próprio método de geração, apesar de não serem muito melhores no caso das

sequências geradas pelo método das adições sucessivas. Note-se no entanto que para os três métodos os menores desvios ocorrem para valores do expoente de Hurst menores, este resultado terá mais a ver com o método de estimação do espectro de potência pois para estes casos devido á anti-persistência das fBf's serão necessários segmentos, para cálculo do periodograma, com menor número de pontos do que no caso de H superiores onde a persistência das fBf's requererá segmentos com mais amostras. Este facto é ilustrado na figura seguinte onde além de uma melhoria do desvio padrão se observa uma melhor estimativa dos expoentes de Hurst elevados quando se utilizam segmentos com 512 pontos.

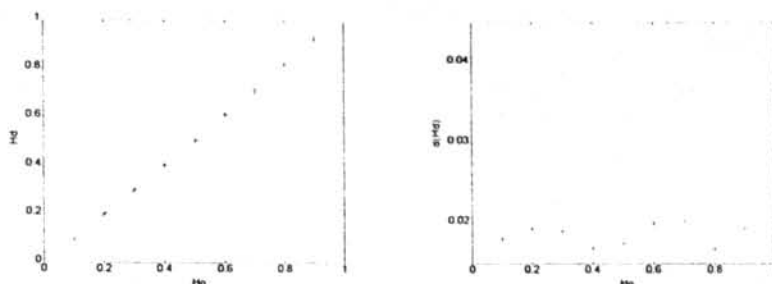


Figura 3-43: Estimação do expoente de Hurst pelos periodogramas com 256 amostras (x) e 512 amostras (+). À direita representa-se o desvio padrão.

3.4.2.2 Método de Hugson (CGSA)

O objectivo deste método, denominado pelos autores (Yamamoto 1993) por *Coarse Graining Spectral Analysis* (CGSA), não é a determinação da dimensão fractal (expoente de Hurst) de uma série temporal, mas a separação das componentes fractais e periódicas presentes num sinal. Este tipo de sinais é bastante frequente em sistemas biológicos, por exemplo o sinal estudado por aqueles autores: frequência de batimento cardíaco em adultos. A variabilidade deste sinal tem sido interpretada por um processo fractal, mas que apresenta componentes periódicas induzidas por exemplo pela respiração. Desta forma é conveniente procurar métodos que permitam a separação das duas componentes evitando a interpretação errada das componentes periódicas.

Este método modeliza as componentes fractais por funções de Brown fraccionárias, utilizando as respectivas propriedades de escalonamento, e utiliza a propriedade dos espectros cruzados de dois sinais sinusoidais com frequências distintas ser nulo devido à ortogonalidade daquelas funções.

Seja $x_c(t)$ um sinal arbitrário contínuo. Fixando uma frequência de amostragem $f_s = 1/T$ definem-se segmentos amostrados do sinal

$$x_l(n, m) = x_c((n + m)T), \quad n = 0, 1, \dots, N-1$$

com N amostras e início no instante mT (variando o valor de m obtêm-se segmentos distintos). Da mesma forma é possível definir segmentos amostrados com frequência f_s/h

$$x_h(n, m) = x_c((hn + m)T).$$

A estimação do espectro de potência, $\bar{S}_{xx}(k)$, e do espectro de potência cruzado, $\bar{S}_{xx_h}(k)$, das sequências amostradas a frequências distintas pode ser efectuada pelo cálculo do periodograma médio e do “periodograma cruzado” médio, tal como anteriormente

$$\begin{aligned}\bar{S}_{xx}(k) &= \frac{1}{N_s} \sum_m S_{xx}(k) = \frac{1}{N_s} \sum_m \frac{1}{L} |X_l(k, m)|^2 \\ \bar{S}_{xx_h}(k) &= \frac{1}{N_s} \sum_m S_{xx_h}(k) = \frac{1}{N_s} \sum_m \frac{1}{L} X_l(k, m) X_h^*(k, m)\end{aligned}, k = 0, 1, \dots, N-1, \quad (3)$$

onde $X_h^*(k, m)$ designa o complexo conjugado da FFT, $X_h(k, m)$, do segmento $x_h(n, m)$, L é um factor de escala já definido e N_s o número de segmentos utilizados.

Considerando que o sinal $x_c(t)$ é uma senoide com frequência f as séries temporais $x_l(n, m)$ e $x_h(n, m)$ serão sinusoides com frequência f e hf , respectivamente. Desta forma e atendendo à ortogonalidade das duas sinusoides o espectro cruzado das duas séries será nulo (na estimação do espectro utilizando o periodograma médio aquele não é, em geral, exactamente nulo devido ao número finito de amostras em cada segmento).

Considerando que o sinal $x_c(t)$ é uma fBf com expoente de Hurst H as séries temporais $x_l(n, m)$ e $h^{-H}x_h(n, m)$ terão a mesma função de distribuição de probabilidade, de acordo com a relação de escalonamento destas funções, pelo que o espectro cruzado das duas séries terá a forma do espectro da série $x_l(n, m)$ a menos de um factor h^{-H} .

Deste dois exemplos pode concluir-se que a amplitude do espectro cruzado de duas séries temporais de um sinal amostradas com frequências distintas é uma estimativa do espectro da componente fractal presente no sinal mais coerente do que o espectro de uma única série. No entanto tal não é totalmente correcto pois se o sinal contem componentes periódicas que formam pares com frequências f e fh o espectro cruzado não será nulo para a frequência f , pois a componente com frequência fh resulta na série $x_h(n, m)$ numa componente com frequência f . Desta forma torna-se impossível a distinção entre este tipo de sinais e sinais fractais que contêm um número “infinito” daqueles pares.

A solução proposta pelos autores do método baseia-se no facto da evolução da fase do periodograma cruzado entre segmentos com atrasos constantes se fase de forma constante quando o sinal é composto apenas por sinais sinusoidais, como se ilustra na Figura 3-44 para as componentes (vectores) não nulas do periodograma cruzado no plano complexo. A situação é bem distinta no caso de sinais fractais, para os quais se observa uma evolução aleatória da fase, como se ilustra na Figura 3-45 (para alguns vectores).

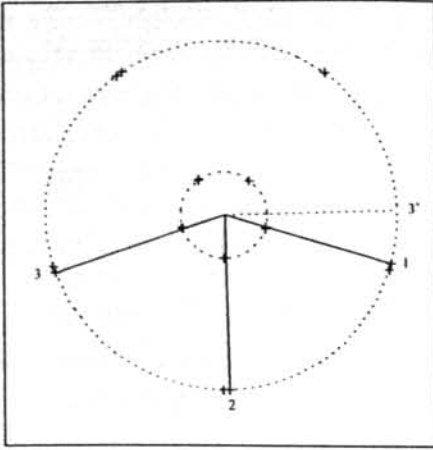
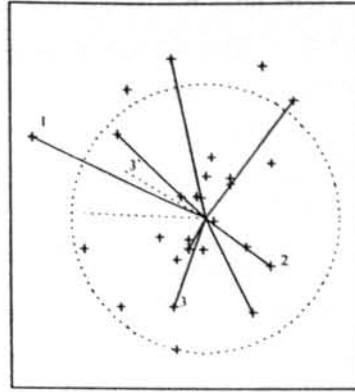
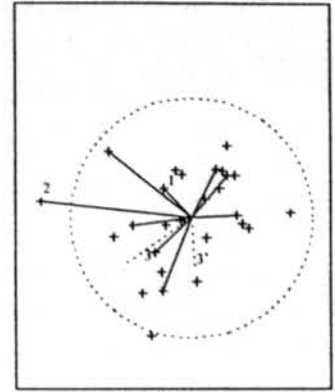


Figura 3-44: Ortogonalização do espectro cruzado de sinal com componentes periódicas.



(a)



(b)

Figura 3-45: Ortogonalização do espectro cruzado de uma função de Brown (fractal).

Representando os vectores de cada periodograma cruzado na forma de Euler

$$S_{xx_h}(k, m) = A(k, m) e^{i\theta(k, m)},$$

onde m designa agora o ordinal do segmento (não o seu início) e o factor h foi omitido por simplicidade, o incremento de fase entre os periodogramas dos segmentos m e $m-1$ será

$$\delta(k, m) = \theta(k, m) - \theta(k, m-1).$$

Supondo que a fase evolui de forma constante então uma estimativa do incremento da fase entre os segmentos m e $m-1$ será o incremento observado nos segmentos anteriores ($m-1$ e $m-2$), ou seja $\hat{\delta}(k, m) = \delta(k, m-1)$. Assim rodando os vectores do periodograma cruzado do segmento m de um ângulo $\pi/2 - \hat{\delta}(k, m)$ obtêm-se vectores

$$S'_{xx_h}(k, m) = A(k, m) e^{i(\theta(k, m) + \pi/2 - \hat{\delta}(k, m))},$$

que são ortogonais aos vectores do periodograma do segmento $m-1$.

Este processo é ilustrado na Figura 3-44 para um vector do periodograma cruzado e três segmentos numerados (1, 2, 3) verificando-se que a ortogonalização do vector 3 relativamente ao vector 2 de que resulta 3' é conseguida devido ao incremento constante da fase do periodograma. Para sinais fractais o resultado do procedimento proposto tanto pode gerar vectores $S'_{xx_h}(k, m)$ colineares a $S_{xx_h}(k, m-1)$ como no caso (a) da Figura 3-45 como vectores perpendiculares como no caso (b) devido ao carácter aleatório evolução da fase do periodograma cruzado.

Tomando a projecção dos vectores $S'_{xx_h}(k, m)$ sobre os vectores $S_{xx_h}(k, m-1)$ tem-se o seguinte resultado

$$A'(k, m) = \frac{|S'_{xx_h}(k, m) \cdot S_{xx_h}(k, m-1)|}{|S_{xx_h}(k, m-1)|} \leq |S'_{xx_h}(k, m)| = A(k, m).$$

Agora, supondo que $A'(k,m)$ tem uma distribuição uniforme no intervalo $[0, A(k,m)]^*$ o espectro de potência da componente fractal de um sinal ($S^F(k)$) será, a menos do factor de escala h^{-4} , estimado pela quantidade

$$S_h^F(k) = \frac{2}{N_s - 2} \sum_{m=2}^{N_s-1} A'(k,m), \quad (4)$$

onde o factor 2 compensa o factor 1/2 do valor esperado de $A'(k,m)$, N_s-2 é o número de segmentos utilizados na média, pois os dois primeiros são utilizados apenas na determinação de $\hat{\alpha}(k,2)$, e se relembra que esta estimativa foi obtida com a séries temporais $x_1(n,m)$ e $x_h(n,m)$. Entretanto o espectro total do sinal ($S^T(k)$) é determinado utilizando os mesmos segmentos no cálculo do periodograma médio, conforme equação (4).

Desconhecendo o valor de H o espectro $S^F(k)$ é determinado calculando os espectros $S_h^F(k)$ e $S_{1/h}^F(k)$ de forma que os factores de escala se anulam no cálculo de

$$S^F(k) = \sqrt{S_h^F(k) \cdot S_{1/h}^F(k)}. \quad (5)$$

O cálculo da percentagem da componente fractal presente num sinal é efectuado de acordo com a expressão

$$\%Fractal = \frac{\sum_{k=1}^{N/2-1} S^F(k)}{\sum_{k=1}^{N/2-1} S^T(k)} \cdot 100, \quad (6)$$

medindo, de acordo com o teorema de Parseval, a percentagem da variância total do sinal que é justificada apenas por um processo fractal.

A implementação do método efectuado no presente trabalho, que tem algumas diferenças em relação ao proposto pelos autores, segue os seguintes passos.

Seja $x(n)$ uma série temporal com N_{data} amostras. Fixando o número de pontos dos segmentos, N , e o espaçamento entre segmentos, l , definem-se N_s segmentos de três séries $x_{1/2}$, x_1 e x_2 com frequências de amostragem 2, 1 e 1/2, respectivamente:

$$\begin{cases} x_{1/2}(n,m) = x(n+ml) \\ x_1(n,m) = x(2n+ml), \\ x_2(n,m) = x(4n+ml) \end{cases} \begin{cases} n = 0, \dots, N-1 \\ m = 0, \dots, N_s-1 \end{cases}$$

Desta forma a série original $x(n)$ é interpretada como a série amostrada a frequência 2, pelo que o domínio de frequências da análise será metade do presente no sinal original. Yamamoto utiliza uma metodologia diferente: duplicando os valores da série original obtém uma série interpolada mantendo assim o domínio útil de frequências.

O valor de N é escolhido como uma potência de 2, para bom desempenho da FFT, l toma o valor $N/2-1$ (ímpar) ou seja existe sobreposição de 50% nos segmentos da série $x_{1/2}$ enquanto

* Esta hipótese é colocada pelos autores do método sem justificação. Seria mais evidente considerar uma distribuição uniforme do ângulo entre os vectores $S'(k,m)$ e $S(k,m)$ sendo neste caso o valor esperado de $A'(k,m)$ igual a $(2/\pi)A(k,m)$. No entanto verificou-se, por experiências, que o factor entre o valor esperado de $A'(k,m)$ e $A(k,m)$ é 1/2 que está de acordo com a hipótese de distribuição uniforme de $A'(k,m)$.

o valor ímpar permite uma maior cobertura da série original pela série x_1 . O número de segmentos é o máximo possível sendo determinado por $N_s = \lfloor (N_{\text{data}} - 3N - l) / (N - l) \rfloor$. A escolha destes parâmetros constitui outra diferença relativamente à implementação de Yamamoto que considera N a maior potência de 2 que não ultrapassa $0.9 \cdot N_{\text{data}}$, fixa N_s igual a 10, determinando l de forma a utilizar o máximo de pontos ($l = \lfloor (N_{\text{data}} - N) / (N_s - 1) \rfloor$). Este conjunto não foi utilizado por o deslocamento dos segmentos l ser muito pequeno resultando que os periodogramas calculados não são suficientemente independentes.

Definidos os segmentos extrai-se-lhes a média e multiplicam-se por uma janela hanning de comprimento N , após o que são calculadas as suas FFT, periodogramas e periodogramas cruzados.

O espectro da componente fractal é determinado utilizando a equação (5) e o espectro total é calculado pelo periodograma médio dos $N_s - 2$ últimos segmentos da série x_1 . A percentagem da componente fractal presente no sinal é estimada pela equação (6).

Os resultados obtidos com este método são apresentados na Figura 3-46 para três séries temporais com 4096 amostras e utilizando segmentos com 256 amostras. Para a primeira amostrada de um sinal com duas componentes sinusoidais (frequências normalizadas à frequência de Nyquist de 0.1 e 0.2) com média nula e variância unitária, observa-se que o método detecta as componentes periódicas e anula o espectro fractal do sinal, que é o objectivo do método. O espectro fractal não é exactamente nulo devido ao número finito de pontos em cada segmento, no entanto note-se que o factor de atenuação das componentes periódicas é da ordem de 10^4 comportando por isso como se fosse nulo.

A segunda série originária de uma função de Brown, igualmente com média nula e variância unitária, gera espectros total e fractal praticamente coincidentes resultando, como seria de esperar, numa percentagem de 100% de componentes fractais.

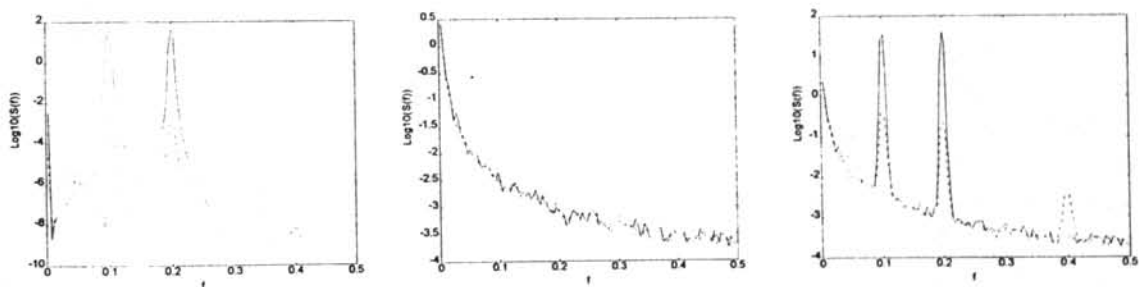


Figura 3-46: Resultado da CGSA aplicada a um sinal com duas componentes sinusoidais, a uma Função de Brown e ao sinal soma dos sinais anteriores, respectivamente da esquerda para a direita. A linha a cheio traça o espectro total do sinal e a linha a tracejado o espectro fractal.

A terceira série é a soma das duas séries anteriores e de acordo com a variância de cada componente será um exemplo de um sinal com 50% de componentes fractais. Os espectros total e fractal obtidos com este método apresentam quatro características, vistas como deficiências, quemerecem ser referidas e discutidas.

A primeira referente à atenuação em grau inferior ao esperado das componentes periódicas presentes no sinal e a segunda referente à ampliação do espectro nas frequências 0.05 e 0.4, formam artefactos que se devem a parcelas do periodograma cruzado que não são ortogonalizadas de forma eficiente pelo presente método, como se explica a seguir. Seja o sinal, x , soma de um sinal com uma componente periódica f , x_f^p , e de um sinal fractal, x^f . Então de acordo com a lineariade da transformada de Fourier a FFT do sinal x será

$$X = X_f^p + X_f^f,$$

e o espectro cruzado de x e do sinal decimado h vezes, x_h , será decomposto nas parcelas

$$\begin{aligned} S_{XX_h} &= X_1 X_h^* = (X_f^p + X_f^f)(X_{hf}^p + X_{hf}^f)^* \\ &= (X_f^p X_{hf}^{p*}) + (X_f^f X_{hf}^{f*}) + (X_f^p X_{hf}^{f*}) + (X_f^f X_{hf}^{p*}) \end{aligned}$$

As duas primeiras parcelas correspondem aos casos puros ilustrados anteriormente e para os quais o método proposto fornece bom resultados. As outras duas originam os artefactos que se observam na Figura 3-47, para os dois casos $h=1/2, 2$; a terceira parcela é responsável pelo artefacto (Tipo I) coincidente com a componente periódica e a quarta pelo outro artefacto (Tipo II).

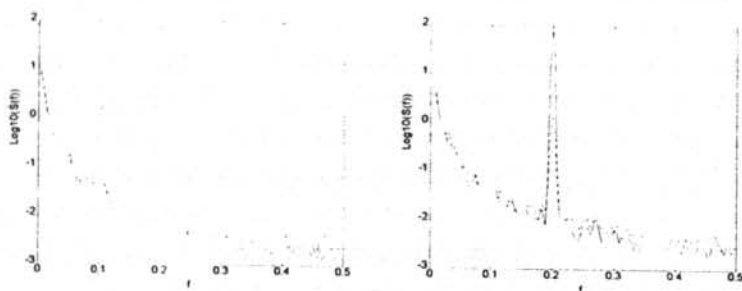


Figura 3-47: Espectros (a cheio) e espectros cruzados (a tracejado) de x_1 e x_h para $h=1/2$ (esquerda) e $h=2$ (direita).

Estes artefactos resultam do facto dos vectores dos periodogramas correspondentes àquelas parcelas manterem o carácter aleatório da fase característico dos sinais fractais pelo que o processo de ortogonalização implementado neste método fracassa. Quando o espectro da componente fractal é calculada pela equação (5) os artefactos do Tipo I mantêm a amplitude enquanto os do Tipo II são atenuados pelo valor correcto do espectro complementar. Para sinais com mais do que uma componente periódica os artefactos nos espectros cruzados são a soma dos artefactos para cada componente, resultando no espectro fractal em artefactos do Tipo I com atenuação (10^2) inferior ao esperado (10^4) e em artefactos do Tipo II com ampliações da ordem de 10^1 - no caso ilustrado na Figura 3-46.

A terceira característica é consequência da não estacionariedade das funções de Brown. Desta forma a variância total do sinal não é estimativa da variância do sinal em cada segmento, inferior à variância total, pelo que a percentagem de componentes fractais calculada a partir da equação (6), fornece valores inferiores ao esperado como se pode concluir pelos resultados da Tabela 3-6. Para 10 sinais independentes com 8192 amostras e as características indicadas nas duas primeiras colunas da tabela, foram calculados os espectros fractal e total utilizando segmentos com 256, 512 e 1024 amostras. Na tabela são apresentados a média e respectivo desvio padrão das percentagens de componentes fractais determinadas pelo método.

Os resultados obtidos nos casos puros estão de acordo com os valores esperados, no entanto nos casos mistos os resultados são sempre inferiores ao valor esperado sendo aproximado quando se aumenta o número de amostras nos segmentos, como seria de esperar pois a variância da componente fractal calculada em segmentos maiores aproxima melhor a variância total daquela componente. Também se verifica uma maior concordância quando o

expoente de Hurst da fBf é menor, facto ser se deve ao carácter anti-persistente destas fBf's como já se referiu anteriormente.

Tabela 3-6: Percentagem de componentes fractais em sinais com características distintas.

H	f	$N = 256$	$N = 512$	$N = 1024$
0.2	-	97 ± 3	94 ± 5	93 ± 7
0.5	-	97 ± 5	97 ± 7	98 ± 18
0.8	-	90 ± 7	94 ± 8	96 ± 16
0.2	0.2	33 ± 1	34 ± 2	40 ± 5
0.5	0.2	13 ± 0.5	20 ± 1	30 ± 5
0.8	0.2	6 ± 0.4	11 ± 2	23 ± 6
-	0.1, 0.2	0 ± 0	0 ± 0	0 ± 0
-	0.01, 0.02	44 ± 1	11 ± 0.2	0.6 ± 0

Os resultados para as percentagem de componentes fractais presentes em sinais apresentados por Yamamoto 1993 concordam em todos os casos com o valor esperado, facto que poderá dever-se ao conjunto de parametros (N , N_s , l) que utiliza, já que o número de amostras em cada segmento é muito elevado e portanto a variância calculada para cada segmento é praticamente igual à variância total do sinal.

A quarta característica, a que se refere a última linha da tabela anterior, tem a ver com resolução espectral do método e a ocorrência de componentes periódicas de baixa frequência. Neste caso o processo de ortogonalização fracassa e o valor da percentagem ultrapassa significativamente o valor esperado, nalguns casos o valor limite 100%.

Face a estes resultados considera-se que esta implementação do método é inconsistente, pelo que será utilizada apenas a título demonstrativo na caracterização das sequências FHR.

3.4.2.3 Método de Higushi

Este método de determinação do expoente de Hurst de fBf's, proposto por Higushi 1988, baseia-se num conceito generalizado de comprimento de uma curva auto-afim e nas propriedades de escalonamento das fBf's.

Dada uma série temporal $x(n)$, $n = 0, 1, \dots, N-1$, o comprimento do gráfico da série $x(n)$ é definido por

$$L = \sum_{n=1}^{N-1} |x(n) - x(n-1)|,$$

agora procurando, tal como no caso de curvas auto-semelhantes, determinar o comprimento da curva utilizando régua padrão de comprimentos distintos, define-se o comprimento do gráfico da série $x(n)$ decimada k vezes por

$$L(k) = \frac{1}{k} \sum_{s=0}^{k-1} \sum_{n=1}^{N(s)-1} |x_k(n, s) - x_k(n-1, s)|, \quad (7)$$

ou seja é a média dos comprimentos para cada uma das k séries distintas (e disjuntas) geradas tomando por origem os instantes $0, \dots, k-1$, identificadas por o índice s e contendo $N(s)$ amostras cada.

Agora de acordo com a propriedade $\Delta V \propto \Delta t^H$ das funções de Brown fraccionárias, se a série $x(n)$ for uma fBf o comprimento das versões decimadas deverá seguir uma lei

$$L(k) \propto \frac{k^H}{k} \propto k^{H-1},$$

onde o termo k^H é o valor esperado da amplitude dos incrementos da série entre intantes separados de $\Delta t = k$ e o termo $1/k$ conta o número de intervalos temporais de amplitude Δt existentes na série original.

Assim observando que uma série temporal tem um comprimento conforme a equação (7) que segue uma lei $L(k) \propto k^\alpha$, para uma gama de valores de k , pode modelizar-se a série por uma fBf com expoente de Hurst $H = \alpha + 1$.

Na implementação do algoritmo utilizaram-se factores de decimação inteiros cujos logaritmos se distribuem de forma uniforme para facilidade de regressão linear dos pontos $(\log(k), \log(L(k)))$ e determinação do expoente α . O número de pontos nos k segmentos da série decimada k vezes considerou-se constante e igual a $N_k = \lfloor N/k \rfloor$, desprezando algumas amostras finais. Após calcular o comprimento da série decimada pela equação (7), o comprimento é corrigido da forma

$$L(k) = \frac{N}{kN_k} L(k),$$

compensando as amostras que não foram utilizadas no cálculo de $L(k)$.

A Figura 3-48 ilustra a aplicação deste método a três realizações de fBf's com $H=0.2, 0.5$ e 0.8 geradas pelo método da síntese espectral. Verifica-se a a função $L(k)$ é bem descrita por uma potência de k e que o valor do expoente estima satisfatoriamente o expoente de Hurst.

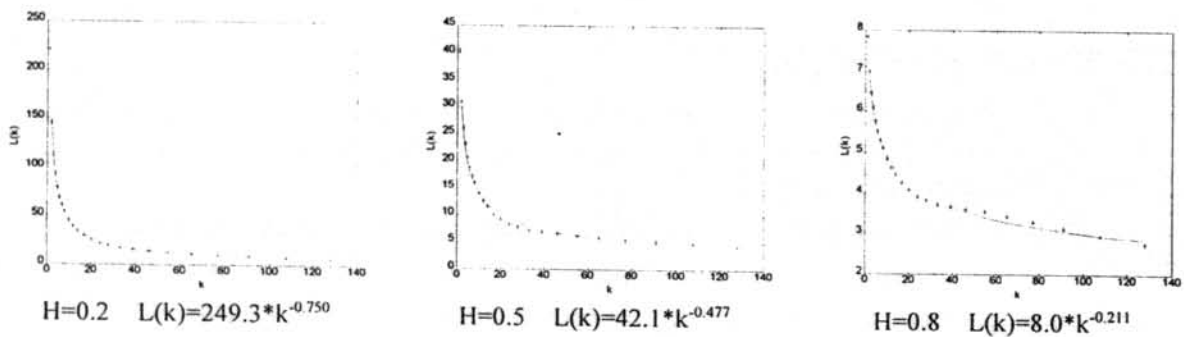


Figura 3-48: Determinação do expoente de Hurst pelo método de Higushi.

Os resultados da figura anterior foram generalizados para outros valores de H e outros métodos de geração. Os resultados da estimação de H , média e desvio padrão para 64 realizações de fBf's para cada caso são apresentados na Figura 3-49.

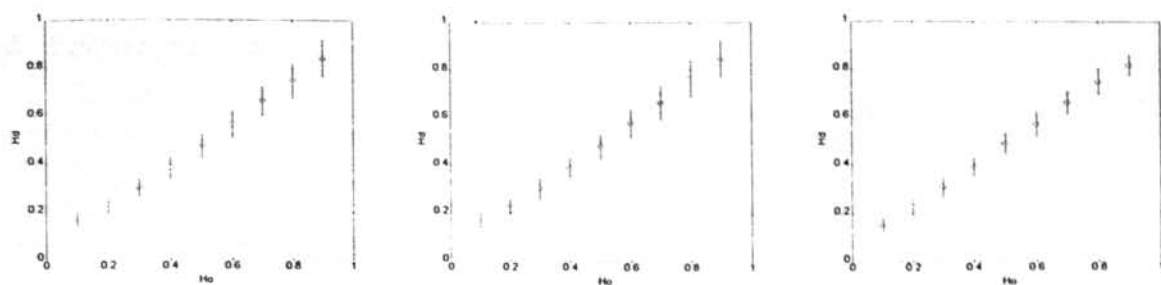


Figura 3-49: Resultados da determinação de H pelo método de Higushi de fBF's geradas por diferentes métodos (mpm, sam e ssm, da esquerda para a direita respectivamente). Média e desvio padrão sobre 64 realizações.

Observa-se que os valores determinados coincidem, em geral e dentro dos erros experimentais, com os valores originais do expoente de Hurst, verificando-se que em todos os casos que os valores determinados se desviam do valor original no sentido de $H=0.5$.

4 Teoria do caos

Este capítulo é dedicado às técnicas baseadas na teoria do caos que foram implementadas no presente trabalho, com o objectivo de caracterizar um sistema dinâmico com base numa série temporal.

Na secção 4.1 definem-se alguns conceitos utilizados neste capítulo e introduzem-se algumas das técnicas de análise de sistemas. Na secção 4.2 descrevem-se os algoritmos implementados.

4.1 Algumas definições

A descrição matemática de muitos fenómenos naturais é efectuada por intermédio de sistemas de equações diferenciais:

$$\frac{d}{dt} X(t) = F(X, t), \quad X(t_0) = X_0.$$

Este sistema de equações descreve a evolução temporal de um vector X de n variáveis de estado x_i de um sistema dinâmico. Este é caracterizado por um campo vectorial $F(X, t)$ que quantifica as interacções entre as variáveis de estado.

A solução deste sistema define, para cada conjunto de condições iniciais (X_0, t_0) , uma curva no espaço de fase ($\mathbb{R}^n$) designada trajectória. Nos casos com interesse prático, estas trajectórias aproximam-se quando t tende para infinito (estado estacionário do sistema) para um conjunto de pontos bem definidos. Este conjunto de pontos denominado atractor caracteriza o comportamento do próprio sistema. Assim um sistema com uma solução constante terá um ponto (de equilíbrio) como atractor, mas se a solução for periódica o atractor será uma curva fechada (ciclo limite) equivalente a um círculo. Quando a solução é uma combinação linear de funções periódicas com frequências incomensuráveis (a solução é quasi-periódica) o atractor toma a forma de um toro, pois apesar de a trajectória poder regressar a um ponto arbitrariamente próximo de qualquer ponto ela não se fecha, ficando por isso sujeita a percorrer uma trajectória sobre um toro.

Além destas formas “simples” para os atractores, estes atractores têm a propriedade de duas trajectórias para condições iniciais arbitrariamente próximas ($X_1(t_0)=X_0$ e $X_2(t_0)=X_0+\delta$, $\delta > 0$) não divergirem, podendo mesmo convergir. Esta propriedade classifica estes sistemas de sistemas preditivos pois as previsões do sistema feitas com base numa observação com erro ϵ do sistema têm erros inferiores a ϵ , dado que as trajectórias se aproximam no futuro.

No entanto em 1963, Lorenz enquanto estudava um modelo muito simples da convecção atmosférica observou, para aquele sistema, que trajectórias arbitrariamente próximas divergiam de forma sistemática, como se pode observar na Figura 4-1 para evolução de uma única variável de estado e na Figura 4-2 para a evolução das trajectórias no atractor do sistema. Este carácter divergentes de trajectórias próximas, também designada por dependência sensível às condições iniciais ou efeito borboleta, dá a este tipo de sistemas um carácter imprevisível ou caótico, pois pouco tempo após uma observação qualquer previsão acerca do sistema não tem qualquer validade.

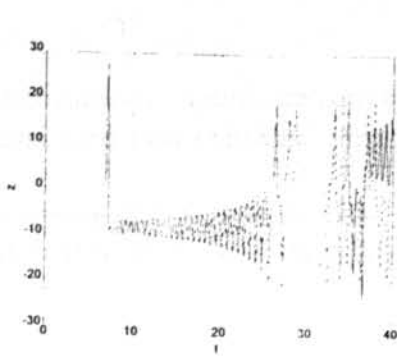


Figura 4-1: Evolução de uma variável de estado, para condições iniciais arbitrariamente próximas.

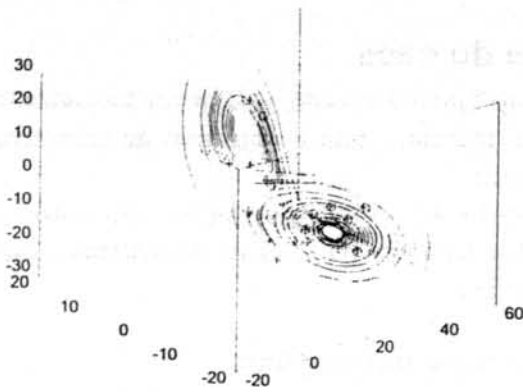


Figura 4-2: Atrator de Lorenz. 'o' e '+' indicam, em instantes de tempo discretos, duas trajetórias com início num ponto arbitrariamente próximo.

Além disso verifica-se que os atratores deste tipo de sistemas, designados atratores estranhos, têm uma estrutura mais complicada que os atratores referidos anteriormente, sendo mesmo fractais porque têm dimensões fraccionárias, a intersecção do atrator com um plano qualquer tem uma estrutura idêntica à dos zeros de uma função de Brown (Figura 3-15).

Esta característica fractal dos atratores estranhos e as dimensões dos outros (zero para os pontos de equilíbrio, um para os ciclos limites e um inteiro maior que um para os toros) permite distinguir os diversos tipos de sistemas e entre os sistemas caóticos a dimensão fractal dos atratores pode estabelecer uma ordem da complexidade do sistema.

Para aplicação destas ideias à caracterização de uma série temporal (no caso sequência FHR) seria necessário definir, primeiro, um modelo do sistema que foi observado, para depois de determinar a trajetória do sistema poder definir o atrator. No entanto existe uma forma mais prática de determinar o atrator de uma sistema observado.

Baseado num teorema provado por Takens em 1980, este método constrói o atrator de um sistema com dimensão fractal D , a partir de uma série temporal $x(n)$ amostrada com frequência f , representando num espaço euclidiano de dimensão $2\lceil D \rceil + 1 = 2N + 1^*$, pontos obtidos da forma

$$[x(i) \quad x(i+1) \quad \dots \quad x(i+2N)].$$

No entanto, a forma geométrica deste objecto, apesar de topologicamente idêntico ao atrator do sistema, depende da frequência de amostragem f . Se a frequência for elevada os pontos são muito correlacionados, sendo oposto o efeito quando a frequência é baixa. Assim é importante a escolha criteriosa da frequência de amostragem, mas quando tal não é possível, como no caso das sequências FHR, os pontos do atrator podem ser obtidos da forma

$$[x(i) \quad x(i+I) \quad \dots \quad x(i+2NI)],$$

sendo o valor do atraso, I , escolhido de forma a que a estrutura do atrator seja evidente, não exista demasiada correlação entre os seus pontos nem uma distribuição quase aleatória dos mesmos.

* $\lceil x \rceil$ designa o menor inteiro maior ou igual a x .

Outro problema deste método de construção do atractor é o facto de necessitar de um parâmetro, D , que se desconhece e cuja determinação é, aliás, o objectivo desta análise. A forma corrente de resolver este ciclo é determinar o atractor para sucessivos valores de N e terminar quando o valor de D , determinado por um qualquer método, não sofrer mais variações. Outra abordagem, implementada no presente trabalho, fixa um valor de N e determina os valores de D_n para todos os atractores em espaços com dimensão 2 até $2N+1$, determinando o valor de D pelo limite dos valores D_i quando N aumenta.

A determinação de D para cada atractor pode ser efectuada quer pela dimensão *box counting* proposta na secção 3.4.1, quer pelos métodos propostos na secção seguinte.

4.2 Dimensão fractal de atractores estranhos

Além da estimação da dimensão fractal de atractores utilizando a dimensão *box counting* foram implementados algoritmos para determinação da dimensão de informação e de correlação, generalizações da primeira. Também foi implementado um algoritmo proposto por Procaccia e Grassberger para determinação da dimensão de correlação.

4.2.1 Dimensão de informação e dimensão de correlação

Estas dimensões fractais são generalizações da dimensão *box counting*, no sentido que a última mede apenas a forma do atractor não fornecendo qualquer informação acerca do sistema dinâmico que gerou o atractor. Uma das abordagens para obter mais informações acerca do sistema consiste em, além de determinar a cobertura do atractor por $N(r)$ caixas de comprimento r , determinar a frequência com que caixa é visitada.

Determinada a probabilidade de ocupação da caixa i , $P(r,i)$, a partir da respectiva frequência de ocupação definem-se duas funções de r :

$$\begin{aligned} H(r) &= - \sum_{i=1}^{N(r)} P(r,i) \log(P(r,i)) \\ C(r) &= \sum_{i=1}^{N(r)} P(r,i) \end{aligned} \quad (1)$$

denominadas entropia e função de correlação, respectivamente. De acordo com Parker 1987, estas funções definem duas dimensões fractais D_I e D_C , da forma daquelas funções*

$$\begin{aligned} H(r) &= H_0 - D_I \log(r) \\ C(r) &\propto r^{D_C} \end{aligned} \quad (2)$$

Estas duas dimensões fractais, juntamente com a dimensão *box counting* D_B , não são contudo independentes. Grassberger 1983b provou que estas dimensões satisfazem a relação

$$D_C \leq D_I \leq D_B, \quad (3)$$

para qualquer atractor, verificando-se, quando a distribuição de pontos no atractor é uniforme, a igualdade das dimensões.

A implementação do algoritmo para determinação das funções entropia e correlação utiliza o algoritmo e estrutura de dados implementados para determinação da dimensão *box*

* A definição proposta por Parker 1987 está errada, sendo a versão correcta a apresentada por Grassberger 1983a

counting, tendo sido acrescentadas funções para cálculo das quantidades da equação (1) a partir da árvore construída.

Os resultados obtidos com este método são apresentados nas Figura 3-38 e Figura 4-4, para os mesmos objectos com que se ilustrou a determinação da dimensão *box counting*.

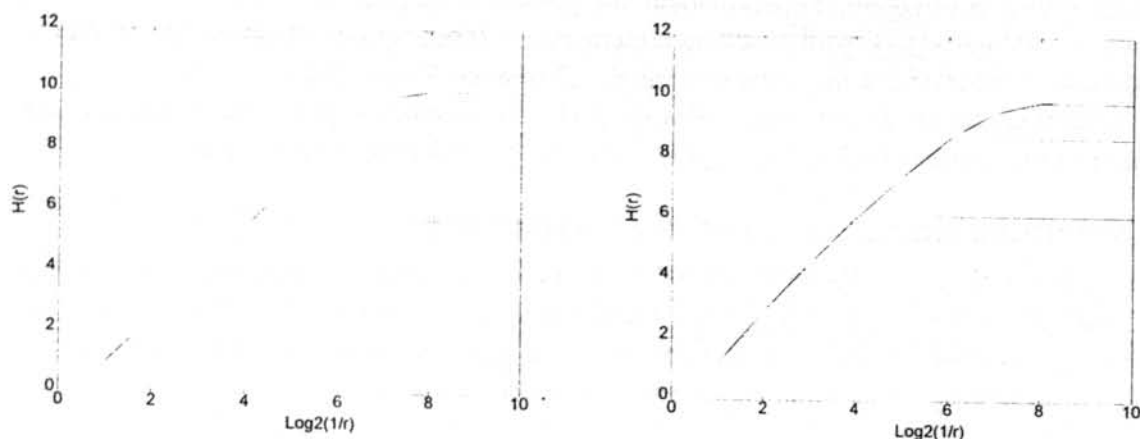


Figura 4-3: Dimensão de informação dos fractais da Figura 3-36. As diferentes (4) curvas correspondem a fractais codificados com 4, 6, 8 e 10 bits. A tracejado indica-se segmento com declive correcto ($D=1.5$).

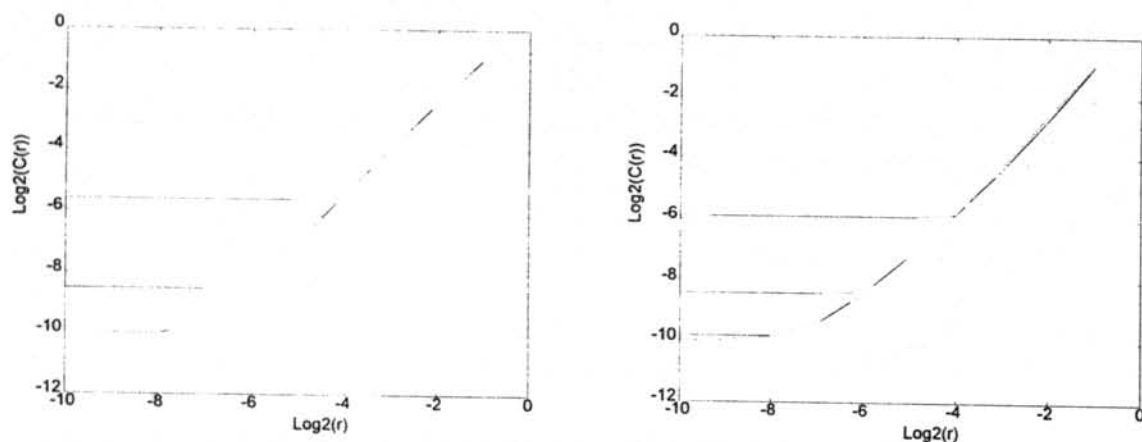


Figura 4-4: Dimensão de correlação dos fractais da Figura 3-36. As diferentes (4) curvas correspondem a fractais codificados com 4, 6, 8 e 10 bits. A tracejado indica-se segmento com declive correcto ($D=1.5$).

Os resultados das dimensões D_B , D_I e D_C obtidos para estes dois objectos fractais são apresentados na Tabela 4-1. Para o fractal geométrico observa-se que estas dimensões fractais satisfazem a relação (3). Para este caso a distribuição de pontos é uniforme no caso de codificação com precisão infinita, no entanto para codificações com número de bits finitos a distribuição é não uniforme, podendo este facto justificar as diferenças observadas. No entanto, no caso do fractal estatístico a relação (3) não se verifica. Entre as dimensões D_I e D_C a relação apenas não se verifica para codificações com um número pequeno de bits, podendo ser justificada por esse facto. Para as dimensões D_B e D_I a diferença poderá dever-se ao número finito de pontos do fractal que não definem a totalidade do objecto, sendo a estimativa do número de caixas inferior ao valor correcto e a estimativa da probabilidade de ocupação deficiente.

Tabela 4-1: Dimensões fractais de dois "atractores".

D=1.5 Codif.	Fractal Geométrico			Fractal Estatístico		
	D_B	D_I	D_C	D_B	D_I	D_C
4 bits	$1.50 \pm .2$	$1.41 \pm .2$	$1.36 \pm .2$	$1.46 \pm .2$	$1.62 \pm .2$	$1.64 \pm .2$
6 bits	$1.54 \pm .06$	$1.50 \pm .06$	$1.47 \pm .06$	$1.40 \pm .06$	$1.49 \pm .06$	$1.54 \pm .06$
8 bits	$1.49 \pm .06$	$1.48 \pm .06$	$1.47 \pm .06$	$1.40 \pm .04$	$1.49 \pm .04$	$1.49 \pm .04$
10 bits	$1.50 \pm .06$	$1.49 \pm .06$	$1.48 \pm .06$	$1.40 \pm .04$	$1.49 \pm .04$	$1.49 \pm .04$

Para o caso da dimensão de correlação, a dificuldade de obter estimativas coerentes da probabilidade de ocupação das caixas pode ser contornada utilizando o algoritmo de estimação da função de correlação proposto por Procaccia e Grassberger de forma a obter estimativas mais realistas da respectiva dimensão.

4.2.2 Dimensão de Procaccia-Grassberger

A definição de função de correlação proposta por Grassberger 1983b é

$$C(r) = \lim_{N \rightarrow \infty} \left(\frac{1}{N^2} \times \# \{ (i, j) : |X(i) - X(j)| < r \} \right),$$

ou seja a função de correlação é a fracção de pares de pontos, $(X(i), X(j))$, do atractor que distam menos que r . Esta definição é equivalente à anterior no entanto o cálculo segundo esta definição é mais fiável que o anterior.

O algoritmo consiste portanto em calcular o histograma, com um número de *bins* adequado, das distância entre todos os pontos de um atractor e calcular o histograma acumulado como estimativa de $C(r)$. Existem no entanto diversas variantes deste algoritmo, conforme a métrica utilizada para cálculo da distância ou a percentagem de pontos utilizados para calcular o histograma.

Na versão implementada utilizou-se a métrica euclidiana, no entanto em vez de se utilizar o valor da distância, utilizou-se o quadrado da distância que além de evitar o cálculo da raiz quadrada também permite definir de forma simples um maior número de *bins* do histograma. Os *bins* utilizados não estão espaçados uniformemente, têm um espaçamento constante numa escala logarítmica de forma a facilitar a determinação da dimensão fractal por regressão linear de $(\log(r), \log(C(r)))$, mas principalmente para aproveitar a representação de números em vírgula flutuante num computador digital. Recordando a representação em vírgula flutuante num sistema binário, a distância r entre dois pontos será expressa da forma

$$r = m \cdot 2^e, \text{ com } 0.5 \leq m < 1,$$

pelo que para *bins* da forma $[2^E, 2^{E+1}[$ a determinação do índice do vector histograma é realizada facilmente determinado o expoente, e , daquela representação. Da mesma forma se x for o quadrado da distância entre dois pontos

$$x = m' \cdot 2^f = m'^2 \cdot 2^{2e} = r^2, \text{ com } f = 2e - (m < 2^{-1/2}) \text{ e } 0.5 \leq m' < 1$$

o número de bins duplica relativamente ao caso anterior. Uma nova duplicação do número do número de bins pode ser efectuada utilizando como índice para o vector histograma utilizando a quantidade

$$g = 2f + (m' \geq 2^{-1/2}).$$

Para evitar a ocorrência de valores de g negativos os pontos do atrator, ou as amostras da série que o constrói, são transformadas de forma que a resolução máxima seja 1. No caso de uma série temporal isto corresponde a tomar uma série de valores inteiro. A relação entre o índice g do vector histograma e os *bins* da distância utilizados será então

$$g = 0 \rightarrow a^{-1} \cdot [0, 1[$$

$$g \rightarrow a^{-1} \cdot [2^{(g-1)/4}, 2^{g/4}[,$$

onde a é o factor de escala usado na transformação dos pontos.

Relativamente à percentagem de pontos utilizados no cálculo do histograma a presente implementação do método calcula o histograma utilizando $N_{\max}$ (20000) pares de pontos, que são os pares formados por todos os pontos do atrator com N_{po} pontos escolhidos aleatoriamente entre os pontos do atrator. Para um atrator com N pontos, o valor de N_{po} é determinado por

$$N_{po} = \min(N, \lfloor N_{\max}/N \rfloor).$$

Além destas características, a presente implementação tem também a particularidade de calcular simultaneamente as funções de correlação para diferentes dimensões do espaço de fase, evitando desta forma o cálculo de repetidas parcelas da distância entre pontos do atrator, já que o valor do atraso usado na construção do atrator é idêntico para todas as dimensões.

O presente método foi aplicado a atractores, em espaços de fase com dimensão desde 2 até 20, de três sistemas (sistema com solução periódica, quasi-periódica (2 frequências) e ao sistema de Lorenz (caótico)), obtidos por construção a partir de uma série temporal codificada com 6.34 bits (80 valores inteiros). Os resultados das funções de correlação obtidos para esses atractores são apresentados na Figura 4-5, verificando-se que seguem satisfatoriamente a lei de potência da equação (2) com expoente adequando (1 para o ciclo limite, 2 para o toro e 2.06 para o atrator estranho). Os casos com maiores desvios do declive correcto ocorrem geralmente para atractores em espaços de fase com dimensão elevada e para atractores com alguma complexidade (toro e atrator estranho) podendo dever-se á menor resolução da estrutura desses atractores.

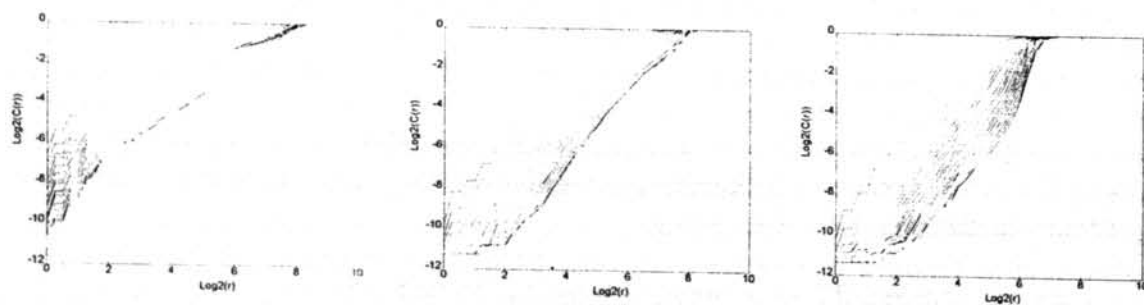


Figura 4-5: Função de correlação para um ciclo limite, um toro e um atrator estranho (atractor de Lorenz). A tracejado segmento com declive D_C correcto.

Além da disparidade entre valores da dimensão real de um atrator e as dimensões estimadas para atratores reconstruídos, em especial para espaços de fase de dimensão elevada - que conduz de acordo com as ideias expostas na página 61 a uma avaliação errada da dimensão do atrator -, também se levanta o problema da determinação da região de valores de r onde é válida a lei de potência para $C(r)$.

Relativamente a este último aspecto, refira-se que para o caso dos atratores das sequências FHR foi implementado um método de pesquisa dessa região, que será descrito posteriormente, mas que não fornece bons resultados para estes atratores. Pois a amplitude da região linear neste caso é significativamente superior à obtida para as sequências FHR e o método de pesquisa é dependente dessa amplitude.

5 Caracterização fractal de sequências FHR

A caracterização fractal de sequências FHR pode ser efectuada de duas formas independentes. A primeira baseada em parte do trabalho de Goldberger 1992 procura modelizar a sequência FHR como uma função de Brown fraccionária (secção 5.1). A segunda baseada nos trabalhos de Chaffin 1991 e Schulman 1994 procura caracterizar aquela sequência determinando as dimensões fractais de atractores obtidos pela representação da sequência num espaço de fase (secção 5.2).

Atendendo à distribuição de casos disponíveis (Tabela 5-1), e ao facto de as duas classes A e FS serem, em sistemas automáticos de classificação, aquelas que apresentam mais dificuldades de classificação e custo mais elevado para o feto, o presente trabalho incidiu, essencialmente, sobre a caracterização de sequências pertencentes a estas classes. Sequências da classe B foram utilizadas como termo de comparação, já que as diferenças entre os traçados de sequências da classe B e das classes A ou FS são evidentes.

Tabela 5-1: Distribuição das sequências pelas diversas classes

Classe	# Sinais	Classe	# Sinais	Classe	# Sinais	Classe	# Sinais
A	35	B	74	C	1	D	13
AD	12	LD	2	DE	7	FS	48
S	3	SH	12	TA	7	BRA	4
MIS ¹	87	SUS ²	48				

¹ Sinais com características de duas ou mais classes. Segmentos com mais de 1000 amostras com características de apenas uma classe foram aproveitados como sinais dessa classe.

² Sinais cuja classificação pelo especialista é duvidosa.

Antes de se aplicarem as técnicas referidas nas secções seguintes, as sequências são processadas para remoção dos artefactos que contenham utilizando o método proposto secção 2.3, não sendo sujeitas a outras técnicas de pré-processamento.

Finalmente, na secção 5.3 são descritas as técnicas de classificação de sequências FHR baseada nas características fractais e apresentados os resultados obtidos.

5.1 Caracterização fractal dos traçados da sequência FHR

Tendo sido enunciada no âmbito de estudos sobre o indivíduo adulto, a hipótese de Goldberger não é válida no caso das sequências FHR, pois, como se verificou na secção 2.2, algumas das classes anormais são tão ou mais complexas que classes normais semelhantes, por exemplo a classe D (Figura 2-5) e AD (Figura 2-6), quando não se identifica a linha base.

Por isso, esta hipótese não constitui a justificação principal para a abordagem tomada neste trabalho; esta é justificada pelo facto de muitas das técnicas de análise de sequências FHR realizarem a análise em múltiplas escalas temporais (duas como proposto por Marques de Sá 1992 e quatro como em Dawes 1981). A observação de que a sequência FHR apresenta estruturas a diversas escalas temporais lembra os objectos fractais, em especial se as respectivas características estatísticas seguirem determinadas leis que permitam estabelecer relações de escalonamento. Procurou-se, por isso, modelizar a sequência por uma função de Brown fraccionária.

O primeiro teste é, de acordo com Goldberger, verificar se o espectro da sequência FHR exhibe uma lei de potência. Nas Figura 5-1 e Figura 5-2 representam-se os espectros de duas sequências FHR, pertencentes às classes B e S, respectivamente. Estes espectros foram obtidos

pelo periodograma dos primeiros 2048 pontos de cada sequência, multiplicados por uma janela hanning, sendo f a frequência normalizada à frequência de Nyquist. Para a sequência da classe S é evidente a sua característica periódica e o espectro da sequência da classe B mostra um espectro de banda mais larga. Esta característica é mais visível na Figura 5-3, onde se representa o espectro, obtido pelo periodograma médio de segmentos com 256 amostras - neste caso a resolução das frequências é menor e mesmo a componente periódica da sequência da classe S não será detectada neste caso, pelo que se perde alguma informação. No entanto para frequências intermédias a região linear do espectro, numa escala bi-logarítmica, define um expoente espectral $\beta=2.14$ a que corresponde um expoente de Hurst $H=0.57$. Para frequência elevadas o espectro torna-se plano devido à presença de ruído branco para frequências superiores a $1/3$, consequência de não se ter efectuado qualquer filtragem da sequência FHR.

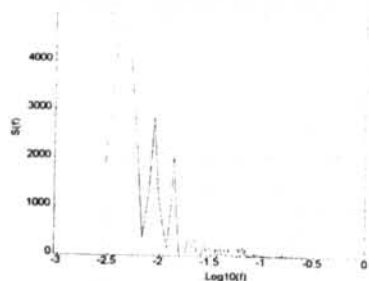


Figura 5-1: Espectro de sequência da classe B.

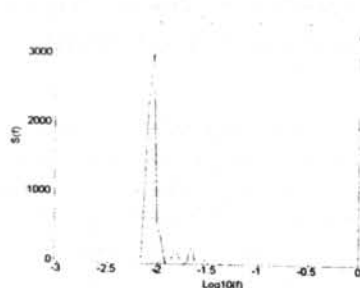


Figura 5-2: Espectro de sequência da classe S.

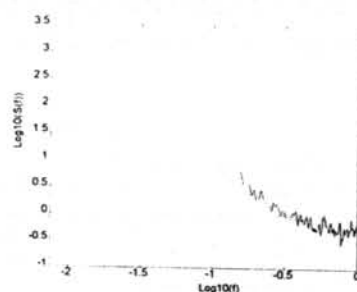


Figura 5-3: Lei de potência do espectro (classe B).

A determinação do expoente de Hurst por esta técnica não é, contudo, muito prática pois exige uma escolha interactiva da região do espectro onde é válida a lei de potência, nem muito fiável devido à existência de componentes periódicas na sequência FHR. Além disso estas séries temporais contêm muito mais informação do que é possível obter com o seu espectro.

O método de Hugson, com o objectivo de separar as componentes periódicas das componentes fractais de um sinal, também não é muito útil para o caso das sequências FHR. As deficiências deste método referidas na secção 3.4.2.2 são ainda mais evidentes no caso das sequências FHR, porque neste caso o número máximo de amostras dos segmentos que se pode utilizar é 256 devido ao número total de amostras do sinal. Assim a maioria das componentes periódicas não serão resolvidas como se verificou no caso dos espectros das figuras anteriores.

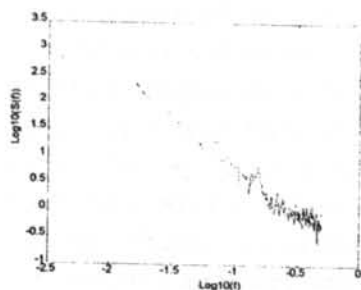


Figura 5-4: Análise CGSA (classe B). Percentagem fractal 74 %.

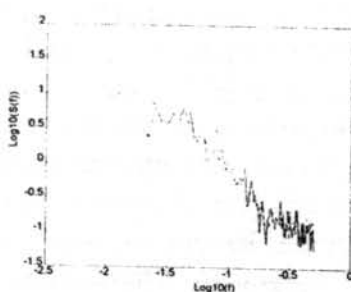


Figura 5-5: Análise CGSA (classe FS). Percentagem fractal 93%.

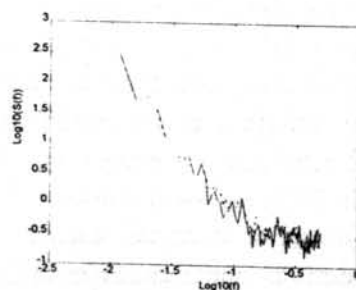


Figura 5-6: Análise CGSA (classe S). Percentagem fractal 64%.

Este facto conduz tal como foi referido aquando da descrição do método à ampliação, para baixas frequências, do espectro fractal como se ilustra na Figura 5-5, apesar de neste caso não ser suficiente para ultrapassar o valor 100%. Nas figuras ao lado os valores da percentagem de componentes fractais também não pode ser questionada quanto ao domínio de validade, verificando-se mesmo para o caso da classe S um valor inferior ao das outras classes, como seria de esperar, mas que não pode ser característico do sinal pelas razões já referidas. Além disso para as três classes, A, B e FS, conclui-se da observação da Figura 5-7 e consulta da Tabela 5-2, que a percentagem fractal não tem qualquer poder discriminativo, um teste estatístico confirma a hipótese de que a média desta característica é idêntica para as três classes. Também são evidentes os requerimentos dos dados para aplicação deste método, o número de amostras no sinal original deve ser no mínimo 2167, para definir pelo menos 10 segmentos por forma a obter estimativas coerentes dos espectros, enquanto o caso da menor percentagem de componentes fractais ($\approx 20\%$) ocorreu para uma sequência com uma correcção deficiente dos artefactos.

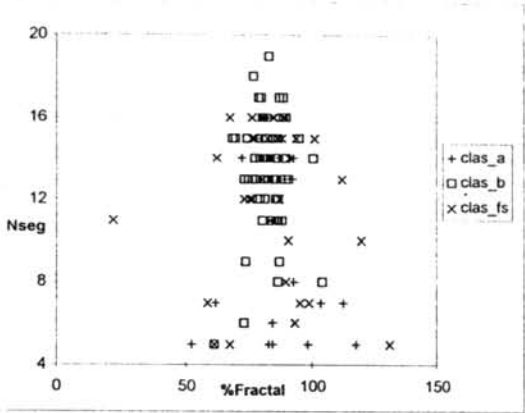


Tabela 5-2: Percentagem de componentes fractais

	A	B	FS
Nseg > 0	87.4±16	82.4±6.8	83.6±6.8
Nseg > 9	85.3±7.5	82.6±5.8	82.6±11 ¹

¹ Neste caso Nseg > 11.

Figura 5-7: Percentagem de componentes fractais em função do número de segmentos utilizados

Os melhores resultados foram obtidos com o método proposto por Higushi 1988, que se ilustra na Figura 5-8, para uma sequência da classe A. Numa escala bi-logarítmica a lei de potência de $L(k)$ é observada como uma recta. No entanto para o caso desta sequência, e para muito outros casos, verificou-se que a função $L(k)$ exibe uma lei de potência mas com dois expoentes distintos para dois conjuntos de valores do factor de decimação k , observando-se por isso duas regiões lineares com declives, eventualmente, distintos nos gráficos em escalas bi-logarítmicas. Também se verificou que os conjuntos de k são semelhantes para quase todas as sequências e que têm fronteira para factores de decimação igual a 10. Este facto permitiu desenvolver um algoritmo que determina automaticamente as melhores regressões lineares no sentido dos mínimos quadrados, nas regiões [3,11] e [10,100] de valores de k . Os valores de k igual a 1 e 2 não foram considerados por serem aqueles que são mais afectados pela presença de ruído na sequência.

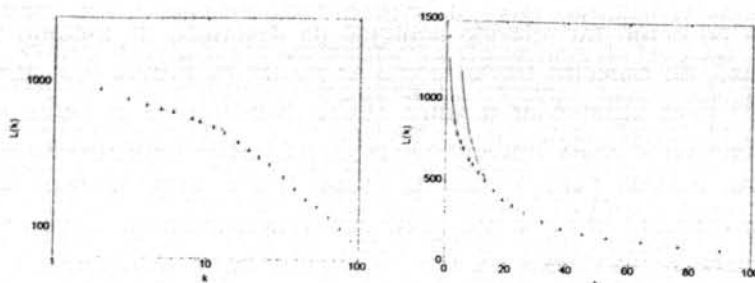


Figura 5-8: Aplicação do método de Higushi a uma sequência da classe A

À semelhança do segundo gráfico da Figura 5-8, nas Figura 5-9 e Figura 5-10 ilustram-se, em escala linear, as aproximações da lei de potência de $L(r)$ para sequências da classe B e FS, exemplificando as hipóteses colocadas acerca da função $L(r)$, que incluem como é óbvio o caso simples de um único expoente (como quase ocorre para a sequência da classe B).

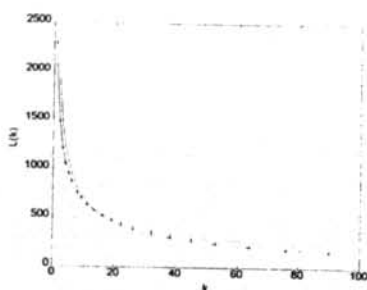


Figura 5-9: Função $L(r)$ para sequência da classe B

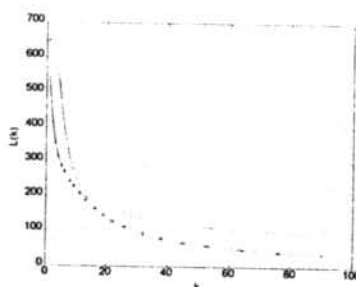


Figura 5-10: Função $L(r)$ para sequência da classe FS

Além do expoente de Hurst, H , determinado a partir do expoente α de $L(r)$, também se determinou o coeficiente independente da recta, C , que será interpretado posteriormente, pois verificou-se ser uma boa característica das sequências FHR. Dado que este último parâmetro é dependente do número de pontos, N , da série em análise, os valores da função $L(k)$ foram normalizadas multiplicando-a pelo factor $2500/N$, correspondendo a tomar sequência com 2500 amostras.

Os valores estimados do expoente de Hurst, H_1 e H_2 respectivamente para k menor e maior que 10, e dos coeficientes C_1 e C_2 , assim como os respectivos intervalos de confiança a 3σ , são apresentados na Figura 5-11 e Figura 5-12.

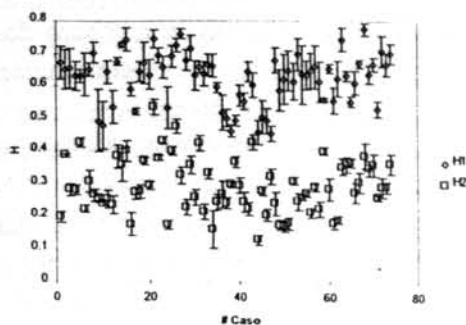


Figura 5-11: Valores de H determinados para sequências da classe B

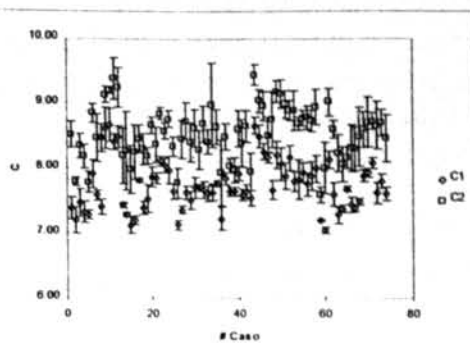


Figura 5-12: Valores de C determinados para sequências da classe B

Com o objectivo de estudar a classificação de sequências das classes A, B e FS baseada nestas características (fractais), representaram-se em *scatter plots* os coeficientes C_1 , C_2 , H_1 e H_2 , sendo exemplos aqueles que são ilustrados na Figura 5-13 e Figura 5-14. No plano H_1 , H_2 observa-se alguma separação entre as classes A e B/FS consequência dos menores valores de H_2 para sequências da classe A. Relativamente às classes B e FS observa-se que ocupam a mesma região do plano, indiciando que a natureza das sequências destas classes será a mesma, diferenciando-se apenas pela menor amplitude das variações como se verifica de seguida.

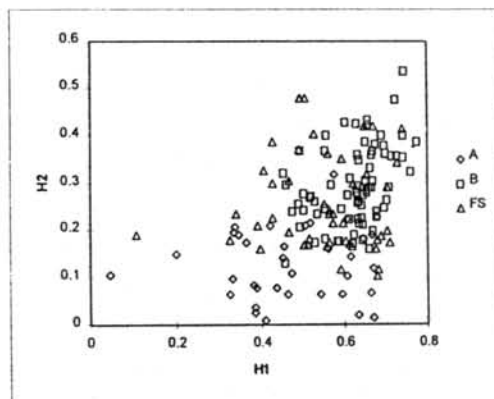


Figura 5-13: Plano de características H_1 e H_2

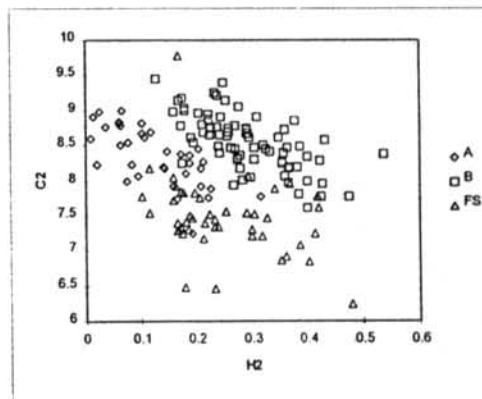


Figura 5-14: Plano de características C_2 e H_2

Notando que o parâmetro C está relacionado com a amplitude do sinal, pois a multiplicação de uma sequência por um factor constante afecta a função $L(r)$ mas apenas no factor de proporcionalidade da potência de k , ou seja de C , é evidente que de acordo com a descrição das classes dada na secção 2.2, este parâmetro tem poder discriminativo daquelas classes, aliás como se pode observar na Figura 5-14. Neste espaço de características (H_2, C_2) a distinção entre as classes B e FS é evidente consequência do menor valor de C para sequências da classe FS (o valor mais elevado pertence contudo à sequência desta classe já identificada aquando da apresentação dos resultados obtidos com o método de Hugson como tendo sido corrigida deficientemente).

5.2 Caracterização fractal da seqüência FHR no espaço de fase

Na caracterização das seqüências FHR utilizando a sua representação num espaço de fase, o primeiro problema que se coloca é o da escolha do valor do atraso para construção dos respectivos atractores de acordo com o método baseado no teorema de Takens.

A solução passa por inspeccionar visualmente os atractores num espaço bidimensional para diversos valores do atraso, como se ilustra na Figura 5-15 para seqüências de 4 classes (A, B, FS e AD), sendo o comprimento dos quadrados que englobam os atractores idêntico para todos eles, 80 bpm, de forma a serem visíveis as consequências da amplitude das seqüências.

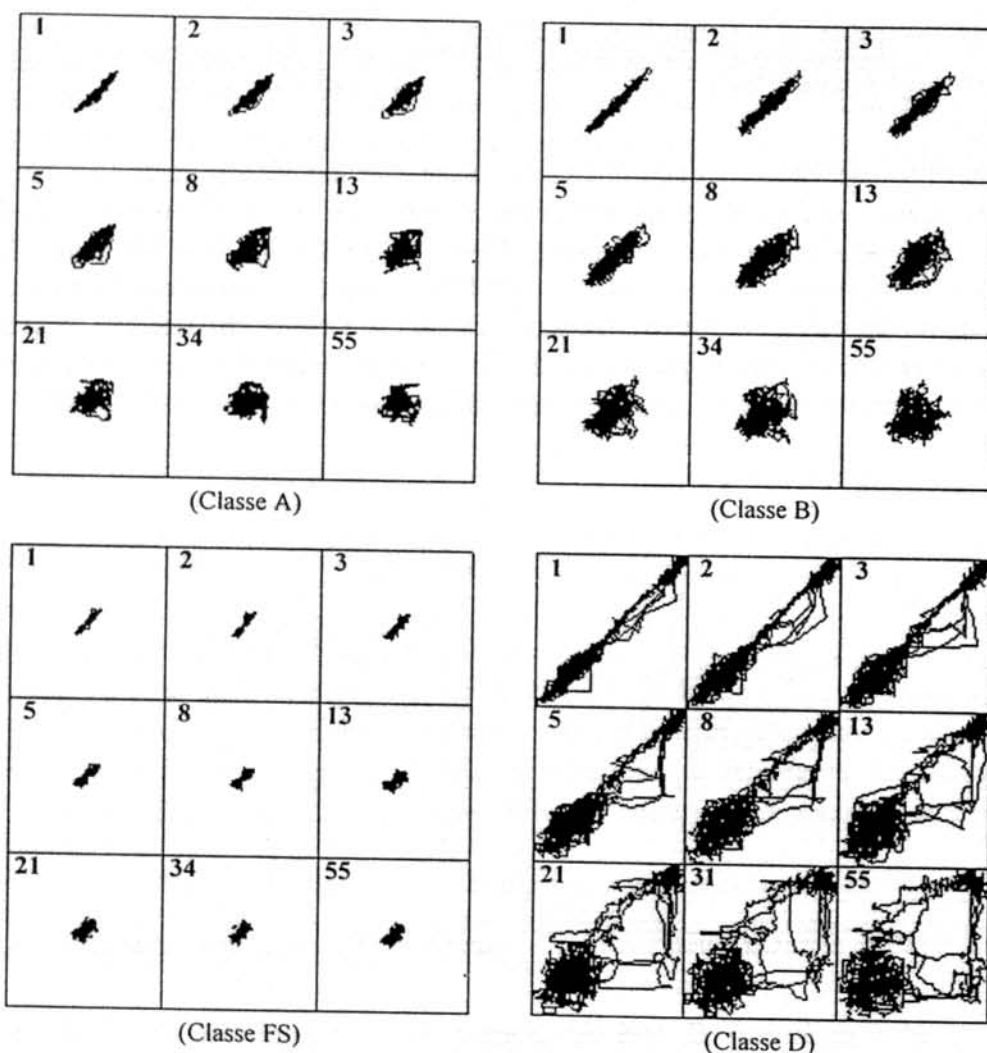


Figura 5-15: Representação de seqüências FHR num espaço de fase bidimensional, para diferentes valores do atraso

Com base nos valores dos atrasos indicados (primeiros números de Fibonacci como proposto por Schulman 1994), considerou-se que o atraso com valor 8 define os atractores com mais informação acerca das seqüências FHR. Além disso extrapolou-se este valor para espaços de fase de dimensão superior, não sendo, contudo, efectuado qualquer teste para esta escolha.

Após a determinação do atraso usado na construção dos atractores foram calculadas as dimensões fractais: Box Counting (definida na secção 3.4.1), dimensão de informação e dimensão de correlação (definidas na secção 4.2.1) e dimensão de Procaccia-Grassberger (PG) (definida na secção 4.2.2), para atractores em espaços de fase com dimensão distintas.

Relativamente às três primeiras ilustram-se, na Figura 5-16, as funções $N(r)$, $H(r)$ e $C(r)$ dos atractores de uma sequência FHR. A dimensão máxima do espaço de fase é 7, pois o número de pontos de que se dispõe é insuficiente para utilizar espaços de dimensão superior, para os quais se torna difícil obter estimativas consistentes das probabilidades utilizadas no cálculo das dimensões de informação e correlação, assim como uma cobertura "fina" do atractor para definição da dimensão *box counting*.

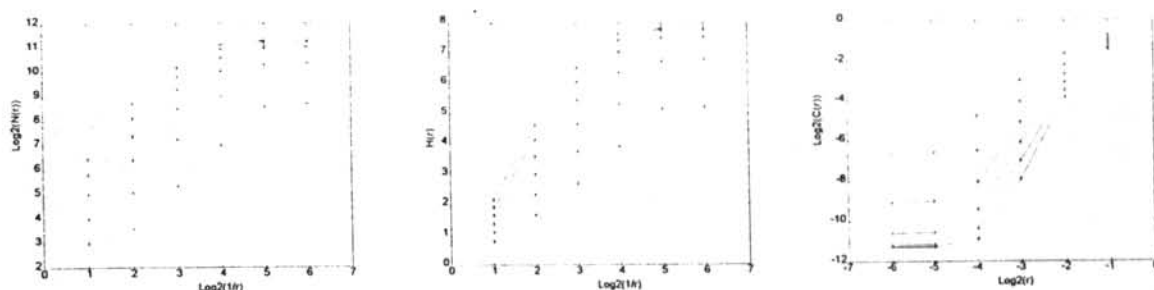


Figura 5-16: Funções $N(r)$, $H(r)$, $C(r)$ para os atractores em espaços de fase com dimensão entre 2 e 7 para uma sequência da classe B. Para dimensões do espaço de fase superiores aumentam os declives das regiões lineares.

As dimensões são estimadas pelos declives das regiões lineares das respectivas funções, que contêm, em geral, um pequeno número de pontos (três ou quatro), devido ao facto das sequências FHR serem sinais de baixa resolução, por exemplo o número de valores distintos das sequências da classe A ou FS raramente ultrapassa os dezasseis. Desta forma os valores dos declives (dimensões) são determinados pelo cálculo do valor máximo do declive da regressão linear de três pontos consecutivos.

Da determinação destas dimensões resultam 18 dimensões ($Dbc2...7$, $Din2...7$ e $Dco2...7$) cuja avaliação como variáveis de caracterização de sequências FHR foi efectuada como no caso de H e C. Na Figura 5-17 apresentam-se as distribuições das sequências para 4 dimensões particulares, observando-se alguma separabilidade das classes FS e A/B, devido principalmente às pequenas dimensões fractais dos atractores das sequências da classe FS, aliás como seria de esperar devido aos valores quase constantes destas sequências.

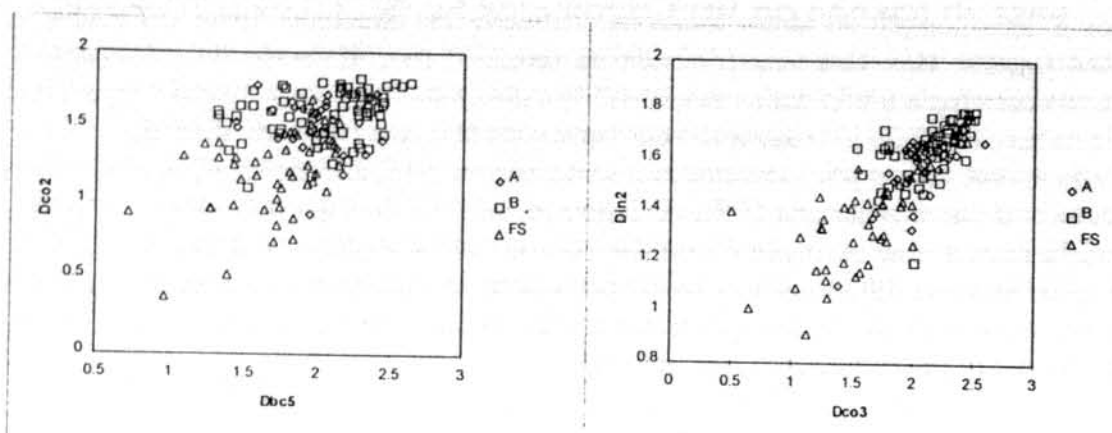


Figura 5-17: Separação das classe por dimensões dos atractores.

Mesmo para outras combinações a separação entre as classes não é tão evidente como no caso de H e C, isto porque estas dimensões fractais são muito correlacionadas e principalmente porque a implementação deste método não tem em conta as medidas lineares dos atractores da Figura 5-15, pois o primeiro passo deste método é encaixar o objecto (atractor) no cubo de comprimento mínimo, não se dando a devida importância àquele factor de escala, apesar de a análise proposta por Schulman 1994 consistir praticamente na medição com uma simples régua dos comprimentos do atractor.

Para o cálculo da dimensão PG utilizou-se a série original, pois a sua resolução máxima já é 1, efectuando o cálculo das funções de correlação para os atractores em espaços de fase com dimensão entre 2 e 20 (é usual na literatura, considerar a dimensão máxima igual a 10, no entanto verificou-se que o eventual valor limite da dimensão PG apenas era atingido para dimensões superiores a 10). As funções de correlação para sequências da classe A, B e FS estão representadas na Figura 5-18, onde são evidentes regiões lineares que definem a dimensão fractal.

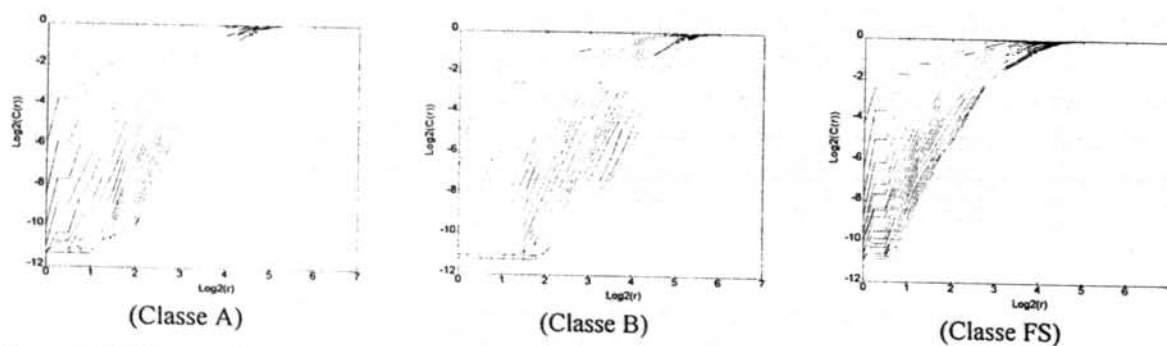


Figura 5-18: Função de correlação $C(r)$, pelo algoritmo de Procaccia-Grassberger. Para dimensões do espaço de fase crescentes o valor da função para $r=1$ diminui.

A determinação da dimensão fractal de cada atractor é realizada por regressão linear da função de correlação de 9 pontos (*bins*) consecutivos, ou seja para intervalos de $\log_2(r)$ com amplitude 2. À dimensão fractal é atribuído, tal como anteriormente, o máximo daqueles declives. Também é registado o valor do ponto médio do respectivo intervalo para

identificação do domínio de aplicabilidade das lei de potência da função de correlação. Na Figura 5-19, representam-se os resultados obtidos para o casos da Figura 5-18.

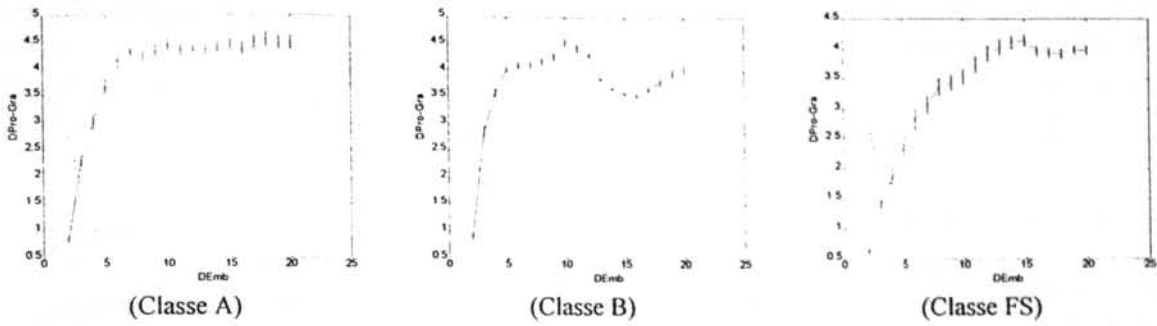


Figura 5-19: Evolução da dimensão PG (linha a cheio, e respectivo desvio padrão) e do ponto médio do intervalo da respectiva região linear de $C(r)$, em função da dimensão DEmb do espaço de fase.

Em vez de determinar o valor limite da dimensão PG, que igualará a dimensão do atrator do sistema, procurou-se aproximar os pontos (DEmb, DPG) pela função

$$DPG = A - Ce^{-B \cdot DEmb},$$

com a estimação dos parâmetros A , B e C efectuada por minimização da função de custo habitual, soma dos quadrados dos desvios, utilizando um método de pesquisa Simplex (função *fmins* do Matlab). Esta função tem a vantagem de englobar casos como a (Classe A) e (Classe FS) da Figura 5-19, assim como casos em que DPG cresce “indefidamente” (tomando B muito pequeno), no entanto não é um modelo óptimo. No caso (Classe B) os parâmetros A , B e C são perturbados pelas oscilações de DPG.

Relativamente à localização da região linear que definiu DPG, são considerados para análise apenas o seu valor médio, MR, e desvio padrão, SR.

5.3 Análise de resultados

Após a determinação das características fractais de sequências FHR como referido nas secções anteriores, o objectivo desta análise é procurar esquemas de classificação automática que determinem a que classe pertence uma sequência com base no seu vector de características fractais.

A classificação bayesiana, um dos esquemas de classificação mais usados, decide a pertença de um objecto descrito por um vector de d características (Marques de Sá 1993)

$$X = [x_1, x_2, \dots, x_d]^T,$$

a uma das k classes de objectos, com base na probabilidade *a posteriori* do objecto pertencer a cada uma das classes c_i :

$$P(c_i | X) = \frac{P(c_i) \cdot p(X | c_i)}{p(X)},$$

onde $P(c_i)$ é a prevalência das classes, ou seja a probabilidade de um objecto não observado pertencer à classe i , e $p(X|c_i)$ é a função densidade de probabilidade (fdp) condicional de X , sendo $p(X) = \sum_{i=1}^k P(c_i) \cdot p(X|c_i)$ a fdp total do vector de características X .

A regra de decisão é*:

$$\left\{ X \in c_i : i = \max_j (P(c_j | X)) \right\},$$

ou seja o objecto é atribuído à classe para a qual é máxima a probabilidade de pertença ($p(c_i|X)$).

Para efectuar o cálculo de $p(c_i|X)$ é necessário formular algumas hipóteses acerca de $P(c_i)$ e $p(X|c_i)$. Relativamente à primeira é habitual considerar que a prevalência das classes são idênticas, ou então são determinadas a partir da frequência de ocorrência de cada classe nos dados. Quanto à segunda considera-se geralmente que as características, X , dos objectos de uma classe particular c_i seguem uma distribuição normal multivariada

$$P(X|c_i) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \exp\left(-\frac{1}{2} (X - M_i)' \Sigma_i^{-1} (X - M_i)\right)$$

com média e matriz de covariância

$$M_i = E[X]$$

$$\Sigma_i = E[(X - M_i)(X - M_i)'].$$

Também é usual considerar que a matriz de covariância é idêntica para todas as classe $\Sigma_i = \Sigma$, obtendo-se desta forma funções discriminantes lineares. Tomando o logaritmo de $p(c_i|X)$ e desprezando valores independentes de X e c_i , a função definida por

$$g_i(X) = \frac{1}{2} ((X - M_i)' \Sigma^{-1} (X - M_i)) - \log(P(c_i))$$

$$= \frac{1}{2} D_M(X, i) - \log(P(c_i))$$

onde $D_M(X, i)$ é designada distância de Mahalanobis, permite determinar a pertença do objecto a uma classe de acordo com a regra

$$\left\{ X \in c_i : i = \min_j (g_j(X)) \right\}.$$

Ou seja, no caso de prevalências iguais, é escolhida a classe para a qual a distância de Mahalanobis do vector de características é menor.

O pacote estatístico, SPSS for Windows 6.1, utilizado nas análise das características usa este tipo de discriminantes.

Supondo que as características fractais determinadas satisfazem as hipóteses sobre $p(X|c_i)$, uma primeira avaliação dos poderes discriminantes da mesma pode ser efectuada por

* O caso mais geral considera que os riscos de cada decisão são distintos e minimiza uma função de risco que é uma combinação linear de todas as probabilidades $p(c_i|X)$.

intermédio de testes de hipóteses sobre a igualdade das médias, rejeitando as características com um nível de significância elevado. Dado que se consideram 3 classes de objectos utilizou-se o teste ANOVA, fornecendo os resultados da Tabela 5-3. Conclui-se que, à excepção das características *A*, *B* e *Din7*, todas elas terão algum poder discriminante daquelas classes, sendo as características *MR*, *SR* e *C2* aquelas que têm médias mais diferentes (pequenos valores de Wilks' Lambda e valores elevados de *F*), devendo ser as que melhor discriminam as classes.

Tabela 5-3: Teste ANOVA

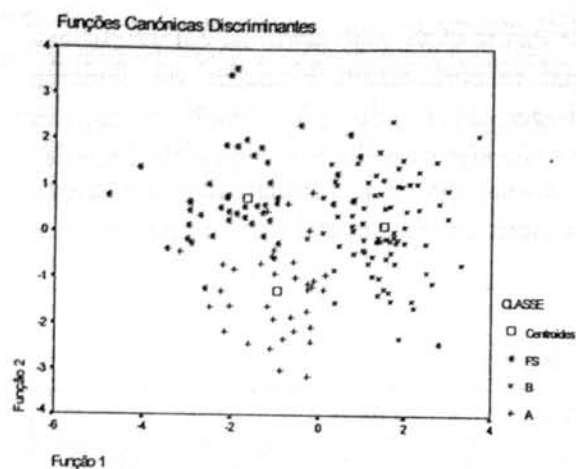
Wilks' Lambda (U-statistic) and univariate F-ratio with 2 and 151 degrees of freedom								
Variable	Wilks' Lambda	F	Signif		Variable	Wilks' Lambda	F	Signif
A	0.97	2.24	0.11					
B	0.92	6.93	0.00		Dco2	0.55	60.67	0.00
C	0.99	0.92	0.40		Dco3	0.56	59.76	0.00
MR	0.35	138.29	0.00		Dco4	0.66	38.76	0.00
SR	0.35	142.72	0.00		Dco5	0.77	22.91	0.00
					Dco6	0.90	8.40	0.00
H1	0.78	21.19	0.00		Dco7	0.95	3.91	0.02
H2	0.63	45.14	0.00					
C1	0.51	71.58	0.00		Din2	0.51	71.57	0.00
C2	0.48	80.87	0.00		Din3	0.62	46.09	0.00
					Din4	0.77	22.37	0.00
Dbc2	0.94	4.75	0.01		Din5	0.89	9.50	0.00
Dbc3	0.86	12.34	0.00		Din6	0.96	3.10	0.05
Dbc4	0.87	11.56	0.00		Din7	0.98	1.64	0.20
Dbc5	0.81	18.01	0.00					
Dbc6	0.87	11.48	0.00					
Dbc7	0.93	5.95	0.00					

A determinação de funções discriminantes foi efectuada em primeira instância para cada grupo de características (*H/C*, *Dbc/Dco/Din* e *DPG*), utilizando a selecção dinâmica de características implementada no comando *discriminant* com os valores por defeito e considerando prevalências idênticas.

Os resultados apresentados são a tabela de resultados, distribuição da classificação determinada pelo discriminante em função da classificação original dos objectos, os coeficientes das funções canónicas discriminantes e a representação dos objectos no plano funções canónicas mais discriminantes.

As funções discriminantes por serem em número de dois, neste problema particular a 3 classes, permitem visualizar, na totalidade, a separação das classes no referido plano. Além disso a linha de separação entre pares de classes é definida pelos pontos do plano equidistantes aos centroides de cada classe.

Tal como já havia sido referido, com base nas Figura 5-13 e Figura 5-14, as características *H* e *C* têm um poder discriminante elevado, tal como se pode comprovar pelos resultados da análise quantitativa. Na Figura 5-20 é evidente a separação entre as três classes, verificando-se que as quatro características determinadas foram utilizadas no cálculo das funções discriminantes.



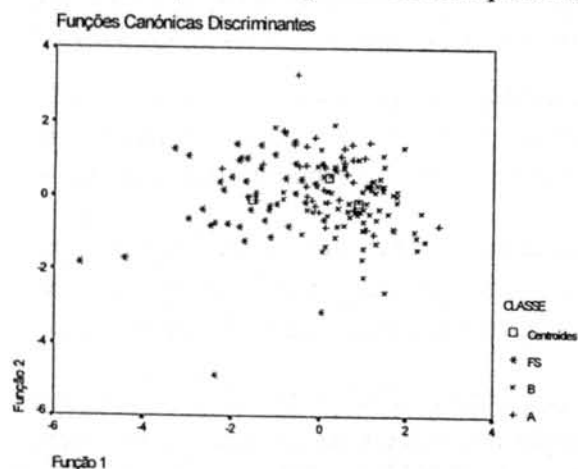
	Func 1	Func 2
C1	-5.059	-6.266
C2	7.545	5.314
H1	-11.511	-11.891
H2	22.511	20.812
(Const)	-23.239	4.525

Classe	Classe Determinada		
	A	B	FS
A	31 (89%)	0 (0%)	4 (11%)
B	3 (4%)	71 (96%)	0 (0%)
FS	3 (6%)	6 (13%)	38 (81%)

Figura 5-20: Discriminantes para as características H e C.

Com era de esperar o número de classificação erradas é bastante baixo (89% de decisões correctas), no entanto ainda existe um risco bastante elevado na avaliação de caso da sequência FS, 3 para A e 6 para B. Como a sequência da classe FS com reconstrução deficiente continua a ser detectada com facilidade - ponto (3,-2.5) - os restantes casos também poderão estar relacionados com reconstruções deficientes mas em menor grau.

Relativamente às características dos atractores, Dbc, Dco e Din, a análise efectuada utilizando os valores dos parâmetro por defeito dos SPSS rejeita todas as características indicando o fraco poder discriminativo destas características. O resultados apresentados correspondem a valores dos parâmetros mínimos. A separação das classes obtida com estas características (Figura 5-21) não é evidente, havendo uma sobreposição da classe A com as classes B e FS, como se pode verificar pela linha A e coluna A da tabela de resultados.



	Func 1	Func 2
DBC6	1.262	-2.296
DBC7	-0.385	4.167
DCO3	1.821	-2.150
DIN2	8.288	11.091
DIN3	-3.872	-6.754
(Const)	-10.247	-1.907

Classe	Classe Determinada		
	A	B	FS
A	18 (51%)	14 (40%)	3 (9%)
B	15 (20%)	55 (74%)	4 (6%)
FS	7 (15%)	2 (5%)	37 (80%)

Figura 5-21: Discriminantes para as características Dbc, Dco e Din.

Quanto às características determinadas a partir de DPG, representadas na Figura 5-22, verifica-se que afinal a separação dos centroides das classes, confirmada pelos resultados da Tabela 5-3, não garante imediatamente a separação das classes, pois a elevada dispersão dos

objectos neste plano provoca, como no caso anterior, uma sobreposição da classe A às classes B e FS, apesar de em menor escala como se comprova pela tabela de resultados.

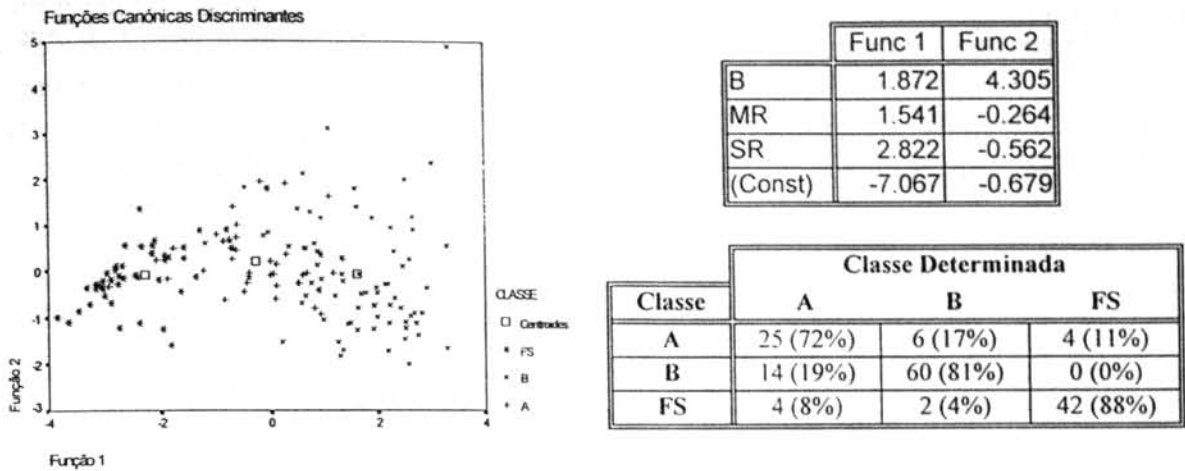


Figura 5-22: Discriminantes para as características DPG.

Os resultados da análise, neste caso, demonstram mais uma vez que as dimensões fractais destes atractores têm um poder discriminante muito inferior às características relacionadas com as medidas lineares do atractor (MR e SR) e, por construção, com a amplitude das sequências FHR. Tomando o parâmetro *A*, do modelo de evolução de DPG com Demb, como o limite de DPG e consequentemente a dimensão fractal do atractor verifica-se, por observação da Figura 5-23, porque esta característica não tem poder discriminante. Esta coincidência da distribuição dos valores de *A* para as três classes pode ter duas justificações. A primeira, referente ao método de determinação da função de correlação pode justificar estes valores por deficiências do método como se observou no caso da determinação da dimensão do atractor de Lorenz. A outra relacionada com um resultado de Chafin 1993, que observou uma grande variabilidade na dimensão fractal dos atractores, poderá justificar igualmente a distribuição de valores observados, em especial a sua dispersão.

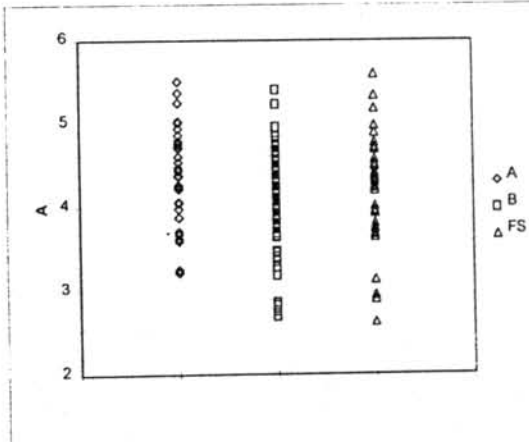


Figura 5-23: Característica *A*.

Relativamente à selecção de características tomando a sua totalidade o resultado mais interessante é a semelhança da distribuição de pontos no plano das funções canónicas neste

caso e no caso da análise das características H e C, revelando mais uma vez a importância destas características. Este facto pode ser confirmado verificando que as três características foram seleccionadas, falta apenas H₁. Relativamente às outras, verifica-se que foi seleccionada apenas uma característica do maior conjunto (Dco3), enquanto das características DPG foram seleccionadas a três que o haviam sido na análise isolada.

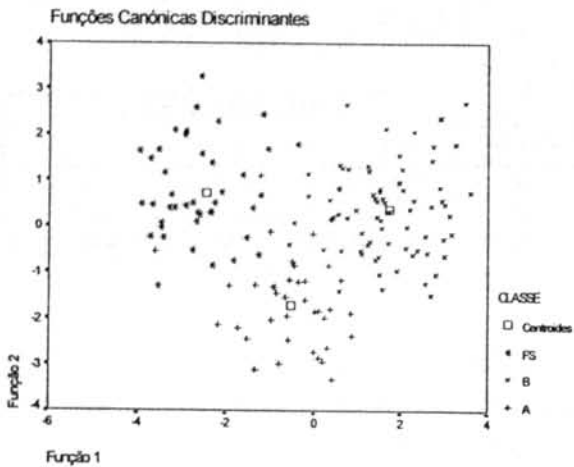


Figura 5-24: Discriminantes para todas as características.

	Func 1	Func 2
B	1.142	-2.073
C1	-0.053	-1.743
C2	0.639	2.128
DCO3	1.164	-0.926
H2	4.381	14.580
MR	1.042	0.206
SR	2.208	-1.803
(Const)	-13.183	-4.796

Classe	Classe Determinada		
	A	B	FS
A	32 (91%)	1 (3%)	3 (8%)
B	3 (4%)	71 (96%)	0 (0%)
FS	2 (5%)	2 (4%)	41 (91%)

No entanto, a percentagem de classificação correcta não aumentou de forma significativa em relação à classificação apenas com H e C, melhorando somente no caso de sequências FS classificadas como sequências B. Além disso, como se pode comprovar pela tabela seguinte, os casos em que ocorrem classificações incorrectas das sequências para os discriminantes baseados nas características (H, C), HIG, baseados nas características DPG e em todas as características, ALL, não são totalmente coincidentes. A maioria dos casos indicados têm uma duração temporal inferior a 20 minutos (obtidos por segmentação de sequências FHR classificadas como mistas), no entanto a duração da sequência não será o factor predominante destas classificações incorrectas, pois existem sequências com duração semelhante que são classificadas correctamente.

Tabela 5-1: Casos com classificação errada.

Caso	Classif.	Discrimin.	Obs. ¹
1	A→FS	ALL	
2	A→FS	HIG	10 min
8	A→FS	HIG / DPG / ALL	15 min
30	A→FS	HIG / ALL	12 min
35	A→FS	HIG	3 min
51	B→A	ALL / HIG	
59	B→A	HIG	
71	B→A	DPG / ALL	4 min
73	B→A	DPG / ALL	15 min
91	B→A	DPG / ALL	

¹ Características da sequência: duração temporal, reconstrução errada, ...

Caso	Classif.	Discrimin.	Obs. ¹
97	B→A	HIG	
113	FS→A	HIG	
121	FS→A	HIG / DPG / ALL	9 min
125	FS→B / FS→A	HIG / DPG	Errada
128	FS→A	DPG / ALL	15 min
139	FS→B / FS→A	HIG / DPG	10 min
141	FS→B	HIG	
144	FS→A	HIG	Classe DE ²
151	FS→B	HIG / DPG	9 min
153	FS→B	HIG	5 min / Errada
155	FS→B	HIG / DPG	7 min / Errada

² Sequência com desaceleração (tipo classe DE).

Mais importante, que a duração, deverá ser a resposta de cada método de determinação das características fractais à presença de artefactos no sinal que analisa. Um exemplo é o caso 125, já referido com tendo um reconstrução deficiente, para o qual os referidos métodos reagem de formas distintas pelo que é classificado, com base nessas características, em classes igualmente distintas. Outro caso é a sequência 144 na qual existe, no intervalo de 20 minutos, uma desaceleração como as observadas em sequências da classe DE. Para este caso, a classificação errada pelo discriminante HIG poderá indicar que o método de Higushi tem capacidade para identificar a desaceleração.

Torna-se, por isso, imperioso que sejam feitos testes que permitam confirmar ou não a sensibilidade dos métodos propostos à presença de ruído, e no caso afirmativo procurar algoritmos de reconstrução de sequências FHR mais eficientes.

6 Conclusões

O método de reconstrução de sequências FHR, apesar das deficiências ilustradas na Figura 2-25, corrige de forma satisfatória a maioria dos artefactos presentes naquelas sequências. A vantagem do método relativamente a outros é a capacidade que possui para extrapolar a existência de batimentos do coração fetal que o monitor não consegue detectar.

No entanto, como sugerido na secção 5.3, a ocorrência de artefactos que subsistem devido às deficiências deste método interfere com os métodos de estimação de características fractais, pelo que é necessário melhorar a qualidade destas sequências: ou na fase de aquisição com monitores fetais mais eficientes ou por melhoria do método proposto. Em relação a esta última solução considera-se que deve basear-se num modelo do sinal cardiotocográfico (as três componentes) como o proposto por Carrillo 1993, de forma a poder obter informação sobre a sequência FHR em casos como o da Figura 2-25.

A caracterização fractal das sequências FHR, definida como a determinação de características fractais e projecto de classificador com base em discriminantes lineares, foi efectuada em dois níveis:

- a) Caracterização fractal dos traçados da sequência FHR;
- b) Caracterização fractal da sequência FHR no espaço de fase.

Relativamente ao primeiro nível, a caracterização consiste essencialmente na determinação do expoente de Hurst, H , de uma função de Brown fraccionária que modelize a sequência. Foram implementados e testados com casos "teóricos" três métodos de estimação daquele expoente. O método baseado no expoente espectral e o método de Higushi determinam dentro dos erros e com um pequeno *bias* que aumenta com H (expoente espectral) ou $|H-0.5|$ (método de Higushi). O outro método (Hugson) que pretendia igualmente determinar a percentagem de componentes fractais num sinal composto não cumpre este objectivo devido à mistura de componentes com caracteres estacionários distintos.

Na aplicação a sequências reais o método do expoente espectral perde qualidades devido à presença de componentes periódicas e à dificuldade de determinar de forma automática a região do espectro que satisfaz a lei de potência. O único método realmente utilizado foi o método de Higushi. Com base neste método obteve-se o resultado mais interessante:

A sequência FHR pode ser considerada um fractal temporal (fBf) com propriedades distintas em pelo menos duas gamas de escalas temporais distintas.

Esta afirmação não é totalmente correcta quando se refere às "escalas temporais", pois este método considera uma série temporal amostrada com frequência constante, que não é o caso da sequência FHR. No entanto, tomando uma frequência média de amostragem (por exemplo a frequência da linha base, com valores entre 110 e 150 bpm, característicos das classes A, B e FS), aquelas gamas de escalas temporais correspondem a $[0, 4 \sim 5.5 \text{ s}]$ e $[4 \sim 5.5 \text{ s}, 40 \sim 55 \text{ s}]$. Com base nestes intervalos, verifica-se que este resultado sobre a sequência FHR é outra forma de exprimir a variabilidade a curto e longo prazo característica daquela sequência.

O classificador obtido com base nestas características, permitiu discriminar de forma eficiente as três classes em estudo.

No outro nível, o objectivo da análise é determinar a dimensão fractal de atractor construído a partir da representação da sequência FHR num espaço de fase. A construção do atractor baseia-se num teorema demonstrado por Takens. As dimensões fractais (dimensão *box counting*, informação e correlação) são estimadas utilizando as coberturas do atractor por caixas alinhadas, sendo a dimensão correlação estimada, também, pelo algoritmo de

Procaccia-Grassberger. As dimensões estimadas desta forma concordam, regra geral, com o valor teórico.

Na aplicação aos casos reais, esta abordagem apesar contar com um número superior de características fractais revelou-se menos produtiva que anterior. Neste caso a discriminação de classes é deficiente e as características mais discriminativas estão associadas aos comprimentos lineares do atractor, ou seja, a forma complexa e o detalhe do atractor traz pouca informação acerca da classe da sequência. Sem pretender, contudo, justificar eventuais falhas nos algoritmos implementados, considera-se, tal como sugere Ruelle 1994, que a aplicação destas técnicas, desenvolvidas para sistemas sob algum controlo, não têm aplicação prática neste tipo de sistemas - abertos, medidos com a precisão possível mas insuficiente.

Face aos resultados apresentados conclui-se que a caracterização fractal de sequências FHR, no domínio temporal, traz algo de novo, apesar das equivalência com conceitos já existentes, ao problema de classificação de sequências FHR. É conveniente por isso os algoritmos propostos sejam sujeitos a testes de sensibilidade ao ruído e, em caso afirmativo, procurar as soluções entre as quais a melhoria da qualidade da sequência FHR, já referida, ou a adaptação destes algoritmos a sinais segmentados. Além disso, a informação obtida a partir das dimensões fractais dos atractores, mesmo que tenha alguma fonte, não se revela fundamental para este problema.

7 Referências bibliográficas

- Carrillo J. F. D., Aguilar J. L. C. & Bernal J. J. P. 1993. Electronic Model as an Aid in Teaching Fetal Heart Rate Monitoring. *Comput. Biol. Med.* 13-3, 151-156.
- Chaffin D. G., Goldberg C. C. & Reed K. L. 1991. The Dimension of Chaos in Fetal Heart Rate. *Am. J. Obstet. et Gynecol.* 165-5, 1425-1429.
- Dawes G. S., Moulden M. & Redman C. W. G. 1990. Limitations of Antenatal Fetal Heart Rate Monitors. *Am. J. Obstet. et Gynecol.* 162-1, 170-173.
- Dawes G. S., Visser G. H. A., Goodman J. D. S. & Redman C. W. G. 1981. Numerical Analysis of Human Fetal Heart Rate: The Quality of Ultrasound Records. *Am. J. Obstet. et Gynecol.* 141-1, 43-52.
- Denton T. A. & Diamond G. A. 1991. Can the Analytic Techniques of Nonlinear Dynamics Distinguish Periodic, Random and Chaotic Signals? *Comput. Biol. Med.* 21-4, 243-263.
- Felgueiras C. S., Marques de Sá J. P. et al 1996. On the Reconstruction of Fetal Heart Rate Sequences. *Actas de Bioeng'96. Sociedade Portuguesa de Engenharia Biomédica.*
- Flandrin P. 1992. Wavelet Analysis and Synthesis of Fractional Brownian Motion. *IEEE Transactions on Information Theory* 38-2, 910-917.
- Fortin C., Kumaresan R. & Ohley W. 1992. Fractal Dimension in the Analysis of Medical Images. *IEEE Engineering in Medicine and Biology*, 65-71.
- Goldberger A. L. 1992. Fractal Mechanisms in the Electrophysiology of the Heart. *IEEE Engineering in Medicine and Biology*, 47-52.
- Grassberger P. & Procaccia I. 1983a. Characterization of Strange Attractors. *Physical Review Letters* 50-5, 346-349.
- Grassberger P. & Procaccia I. 1983b. Measuring the Strangeness of Strange Attractors. *Physica D* 9D, 189-208.
- Higushi T. 1988. Approach to an Irregular Time Series on the Basis of the Fractal Theory. *Physica D* 31, 277-283.
- Higushi T. 1990. Relationship Between the Fractal Dimension and the Power Law Index for a Time Series: A Numerical Investigation. *Physica D* 46, 254-264.
- Kobayashi M. & Musha T. 1982. 1/f Fluctuation of Heartbeat Period. *IEEE Transactions on Biomedical Engineering BME* 29-6, 456-457.
- Mandelbrot B. B. & van Ness J. W. 1968. Fractional Brownian Motions, Fractional Noises and Applications. *SIAM Review* 10-4, 422-437.
- Mandelbrot B. B. 1982. *The Fractal Geometry of Nature*. W. H. Freeman.
- Marques de Sá J. P., Moura A. C. et al 1992. Análise Automática de Sinais cardiográficos. *Acta Médica Portuguesa* 5-1, 28-30.
- Marques de Sá J. P. 1991. Análise de Dados. DEEC-FEUP.
- Parker T. S. & Chua L. O. 1987. Chaos: A Tutorial for Engineers. *Proceedings IEEE* 75, 982-1008.
- Peitgen H.-O. & Saupe D. 1988. *The Science of Fractal Images*. Editors. Springer-Verlag.
- Ruelle D. 1994. O Acaso e o Caos. Relógio d'Água.
- Sande C. H. & Gomez A. H. 1981. Aid to Diagnosis of Foetal Bradycardias Using the Autocorrelation Function. *IEE Proceedings* 128A-8, 571-576.
- Saupe D. 1988a. Algorithms for Random Fractals. in *The Science of Fractal Images*. Springer-Verlag.
- Saupe D. 1988b. A Unified Approach to Fractal Curves and Plants. in *The Science of Fractal Images*. Springer-Verlag.
- Schulman H., Schulman S. et al 1994. Antepartum Computerized Fetal Heart Rate, Chaos Analysis, Blood Gases and Perinatal Outcome. *Journal of Maternal-Fetal Investigation* 4, 59-63.
- Stewart I. 1995. Os Problemas da Matemática, Gradiva.
- Tuck D. L. 1982. Improvement in Doppler Ultrasound Human Foetal Heart Rate Records by Signal Correlation. *Medical & Biological Engineering & Computing* 20, 357-360.
- Voss R. F. 1988. Fractals in Nature: From Characterization to Simulation. in *The Science of Fractal Images*. Springer-Verlag.
- Yamamoto Y. & Hughson R. L. 1993. Extracting Fractal Components from Time Series. *Physica D* 68, 250-264.

8 Apêndices

8.1 Programas de reconstrução da sequência FHR

<pre>function [x,y,z,ndes]=corri1(filename) %CORRIGE Corrige os artefactos de uma sequencia FHR % [x,y,z,ndes]=corri1(filename) le o ficheiro filename % que contem o sinal carditocografico e retorna a % sequencia FHR em x, a sequencia corrigida em y e a % sequencia interpolada em z. ndes numero de artefactos % em cada passo do processo de correcção desvio= 10; % Amplitude maxima das descontinuidades ndes=zeros(1,5); % Conta artefactos em cada fase x=condinit(filename); % Condições iniciais [x,y,b,e]=clas(x,desvio); % Determinação dos artefactos if length(b) ndes(1)=length(b); [b,e]=adjcont(y,b,e,desvio); % Correcao de descontinuidades " verdadeiras" [y,b,e]=quebras(x,b,e,y,desvio); % Correcao de artefactos com "continuidade" entre os extremos [y,b,e]=contind(x,b,e,y,desvio); if length(b) ndes(2)=length(b); [y,b,e]=contind(x,b,e,y,2*desvio);</pre>	<pre> if length(b) % Correcao de limites de artefactos ndes(3)=length(b); [b,e]=joindesc(x,b,e,8,desvio); [y,b,e]=contind(x,b,e,y,desvio); [b,e]=joindesc(x,b,e,16,desvio); [y,b,e]=contind(x,b,e,y,desvio); ndes(4)=length(b); [b,e]=joindesc(x,b,e,30,2*desvio); [y,b,e]=contind(x,b,e,y,2*desvio); [y,b,e]=corrdes(y,b,e); ndes(5)=length(b); end end end [x,y]=undoinit(x,y); [z,x,y,q]=calcamos(x,y); y=reconst1(z,q); y=round(y);</pre>
---	--

<pre>function x=condinit(filename) %CONDINIT Le e trata condicoes extremas % de um sinal. % x=condinit(filename) le o ficheiro filename % que contem o sinal carditocografico e retorna % a sequencia FHR com tratamento das % condicoes extremas do sinal. % Introducao de segmentos iniciais e finais com % valores estimados constantes. global NPointIniFim % numero de pontos nos % segmentos a introduzir NPointIniFim=6; if isstr(filename) % Le ficheiro x=leitura(filename); if x == []; return; end x=x(:,1);</pre>	<pre>else % Ja e o sinal [n,m]=size(filename); if (n > m) x=filename(:,1); else x=(filename(1,:))'; % x e um vector coluna end end % Introducao dos segmentos - maximos dos % histogramas em segmentos com 50 pontos bin=40:200; range=1:50; h=hist(x(range),bin); [b,i]=max(h); x=[bin(i)*ones(NPointIniFim,1); x]; h=hist(x(length(x)-range),bin); [b,i]=max(h); x=[x; bin(i)*ones(NPointIniFim,1)];</pre>
--	---

```

function [x,y,b,e,desvio]=clas(x,desvio)
%CLAS Classifica os pontos de uma sequencia FHR
% [x,y,b,e,desvio]=clas(x,desvio) procura
% descontinuidades na sequencia original x,
% com amplitude superior, em valor absoluto,
% a desvio e retorna os extremos de sequencias %
% contiguas de descontinuidades ( b e e).
% Assim como uma estimativa do sequencia
% corrigida ( y = x).

dax=abs(diff(x));

ind=find(dax > desvio);

y=zeros(size(x));
% marca o ponto onde ocorreu o salto
y(ind)=ones(length(ind),1);

% e o ponto anterior
y(ind+1)=ones(length(ind),1);

[ind1,ind2]=detreg(y); % Regioes nao validas

b=ind1;
e=ind2;

y=x;

```

```

function [b,e]=adjcont(x,b,e,desvio)
%ADJCONT Procura do maior grau de
% continuidade entre extremos de
% uma descontinuidade
% [b,e]=adjcont(x,b,e,desvio), recebe
% x: vector de dados original
% b: vector de inicios de descontinuidades
% e: vector de finais de descontinuidades
% desvio: valor do incremento maximo, para
% continuidade e corrige as descontinuidades,
% cujos extremos diferem menos que desvio,
% retornando os vectores b e e actualizados

% procura de sequencia que sao mais
% "contínuas" considerando os pontos vizinhos

da=[-1 0 1 0 -1];
dp=[ 0 1 0 -1 1];
nseq=[2 2 2 2 3];

dextr=(x(e)-x(b));

% Sequencias com mais de dois pontos, com
% descontinuidades superiores a desvio, mas
% inferiores a 3xdesvio (desaceleracoes) ou
% 2xdesvio (aceleracoes)

ind=find(dextr >= -3*desvio & dextr <= 2*desvio &
abs(dextr) > desvio & (e-b) > 2);

if (length(ind) == 0) return; end

for k=1:length(da)
d(:,k)=x(e(ind)+dp(k)) - x(b(ind)+da(k));
end

[dmin,imin]=min(abs(d'));
ind2=find(dmin <= desvio);

for k=1:length(ind2)
k2=ind2(k);
b(ind(k2))=b(ind(k2))+da(imin(k2));
e(ind(k2))=e(ind(k2))+dp(imin(k2));
end

```

```

function [y,b,e]=quebras(x,b,e,y,desvio)
%QUEBRAS Detecta e corrige descontinuidades
% com comprimento 2
% [y,b,e]=quebras(x,b,e,y,desvio), recebe
% x: vector de dados original
% b: vector de inicios de descontinuidades
% e: vector de finais de descontinuidades
% y: vector de dados interpolado
% desvio: valor do incremento maximo, para
% continuidade e corrige as descontinuidades
% com comprimento 2 retornando os vectores y,
% b e e actualizados

x=x(:);

da=[-1 0 -1];
dp=[0 1 1];
nseq=[2 2 3];
% Sequencias com dois pontos
ind=find((e-b)==1);
if (length(ind) == 0)
    return
end

for k=1:3
    d(:,k)=(x(e(ind) + dp(k)) - x(b(ind) + da(k))) /
        sqrt(nseq(k));
end

[dmin,imin]=min(abs(d'));
ind2=find(dmin <= desvio);

for k=1:length(ind2)
    ind3=(b(ind(ind2(k))) + da(imin(ind2(k)))) :
        (e(ind(ind2(k))) + dp(imin(ind2(k))));
    y(ind3)=x(ind3(1)) + (0:(length(ind3)-1))*
        (x(ind3(length(ind3)))-x(ind3(1))) /
        (length(ind3)-1);
end

e(ind(ind2))=[]; % Actualiza valores das
b(ind(ind2))=[]; % descontinuidades
d=[];

```

```

function [y,b,e]=contind(x,b,e,y,desvio)
%CONTIND Detecta e corrige descontinuidades cujos
% extremos diferem menos que desvio
% [y,b,e]=contind(x,b,e,y,desvio), recebe
% x: vector de dados original
% b: vector de inicios de descontinuidades
% e: vector de finais de descontinuidades
% y: vector de dados interpolado
% desvio: valor do incremento maximo, para
% continuidade e corrige as descontinuidades, cujos
% extremos diferem menos que desvio,
% retornando os vectores y, b e e actualizados

% Sequencias cujos extremos diferem menos que desvio
ind=find(abs(x(e)-x(b)) <= desvio);

if (length(ind) == 0)
    return;
end

y2=y;

for k=1:length(ind)
    xi=[b(ind(k)) e(ind(k))];
    ind2=(xi(1)+1):(xi(2)-1);
    y2(ind2)=interp1(xi,y(xi),ind2);
end

y=y2;
e(ind)=[];
b(ind)=[];

```

```
function [b,e]=joindesc(x,b,e,ampl,desvio)
%JOINDESC Junta discontinuidades que nao
% distam mais de ampl
% [b,e]=joindesc(x,b,e,ampl,desvio), recebe
% x: vector de dados
% b: vector de inicios de discontinuidades
% e: vector de finais de discontinuidades
% ampl: valor da distancia maxima entre
% discontinuidades.
% desvio: valor maximo das discontinuidades
```

```
ind=1:length(e)-1;
if (length(ind) == 0)
return;
end
```

```
% Distancia entre discontinuidades consecutivas
dinter=b - [-100; e(ind)];
% Amplitude do desvio entre aquelas
xinter=abs(x(b)-[-100; x(e(ind))]);
```

```
% Agrupamentos dos artefactos que nao
% distam mais de ampl amostras
clasgrup=cumsum(~(dinter <= ampl & xinter <=
desvio));
```

```
% Para todos os grupos formados
for k=1:max(clasgrup)
ind=find(clasgrup == k);
lind=length(ind);
% Grupos com mais de um elemento
% Portanto agrupa discontinuidades
if (lind > 1)
e(ind(1:(lind-1)))=zeros(lind-1,1);
b(ind(2:lind))=zeros(lind-1,1);
end
end
```

```
ind=find( e == 0);
e(ind)=[];
ind=find( b == 0);
b(ind)=[];
```

```
function [y,b,e]=corrdes(x,b,e)
%CORRDES Detecta e corrige discontinuidades
% complexas
% [y,b,e]=corrdes(x,b,e), recebe
% x: vector de dados original
% b: vector de inicios de discontinuidades
% e: vector de finais de discontinuidades
% retornando os vectores y, b e e actualizados
```

```
y=x;
```

```
for k=1:length(b)
ind1=b(k):e(k);
ind2=(b(k)-5):(e(k)+5);
ind3=5+(1:length(ind1));
z=fmeanrank(x(ind2),9,5); % Filtro
y(ind1)=z(ind3);
end
```

```
function [x,y]=undoinit(x,y)
%UNDOINIT Funcao utilizada na reconstrucao
% de uma sequencia FHR
% [x,y]=undoinit(x,y) desfaz as condicoes
% iniciais introduzidas em condinit
```

```
global NPointIniFim
```

```
ind=(NPointIniFim+1):(length(x)-NPointIniFim);
y=y(ind);
x=x(ind);
```

```
function [z,x,y,q]=calcamos(x,y)
%CALCAMOS Funcao utilizada na reconstrucao
%          de uma sequencia FHR
%          [z,x,y,q]=calcamos(x,y). Dado o sinal
%          original x e o sinal estimativa y
%          corrige-se y baseado no quociente y/x.
%          Retorna o sinal corrigido z e
%          os quocientes (inteiros) q
```

```
q=y./x;
```

```
% Pontos correctos
ind1=find(abs(q-1) < 0.05);
% Falha de uma deteccao
ind2=find(abs(q-2) < 0.1);
% Falha de duas deteccao
ind3=find(abs(q-3) < 0.15);
```

```
y(ind2)=x(ind2);
y(ind2)=2*x(ind2);
y(ind3)=3*x(ind3);
```

```
q(ind2)=2*ones(size(ind2));
q(ind3)=3*ones(size(ind3));
```

```
clear ind2 ind3
```

```
% Valores nao validos que restam
```

```
ind=find (q-fix(q));
```

```
m=zeros(size(q));
```

```
% Marca valores nao validos
```

```
m(ind)=ones(size(ind));
```

```
% Identifica regioes nao validas em m
```

```
[ini,fim]=detreg(m);
```

```
s=0;
```

```
% Corrige as sequencias identificadas
```

```
for k=1:length(ini)
```

```
    ind=ini(k):fim(k);
```

```
% Soma do quociente para as amostras de uma
```

```
% descontidade
```

```
    s=s+sum(q(ind));
```

```
    s2=fix(s+0.5);
```

```
    s=s-s2;
```

```
% Sinal "menos" para distinguir dos casos
anteriores
```

```
    q(ind)=-s2*ones(size(ind));
```

```
end
```

```
z=y;
```

```
function y=reconstl(x,q)
%RECONST1 Funcao utilizada na reconstrucao
%          de sequencias de FHR
%          y=reconstl(x,q). Forma o vector y a partir
%          dos vectores x e q
%          Se q == 1, y=x
%          Se q == 2, y=[x x]
%          ...
```

```
x=x(:);
q=q(:);
dq=diff(q);
% Transicoes de valores de q
ind=find(dq)+1;
% Valores iniciais e finais
ind=[1; ind; length(x)+1];
```

```
y=[];
```

```
for k=1:length(ind)-1
```

```
    if ((q(ind(k))) == 1)
```

```
        y=[y; x(ind(k):(ind(k+1)-1))];
```

```
    elseif (q(ind(k)) > 0)
```

```
        for l=ind(k):ind(k+1)-1;
```

```
            y=[y; x(l)*ones(q(l),1)];
```

```
        end
```

```
    elseif (q(ind(k)) < 0)
```

```
        nam=abs(q(ind(k)))-1; % 0:nam
```

```
        ind1=ind(k):(ind(k+1)-1);
```

```
        ind2=0:(length(ind1)-1);
```

```
        if nam
```

```
            ind3=(0:nam)*(length(ind1)-1)/nam;
```

```
        else
```

```
            ind3=mean(ind2);
```

```
        end
```

```
        if (length(ind1) > 1)
```

```
            y=[y; interp1(ind2,x(ind1),ind3)];
```

```
        else
```

```
            y=[y; x(ind1)*ones(nam+1,1)];
```

```
        end
```

```
    end
```

```
end
```

```
function [ini,fim]=detreg(x)
%DETREG Detecta regioes num sinal binario
% [ini,fim]=detreg(x) detecta regioes do sinal
% binario x com valor 1.
% ini contem os indices correspondentes ao
% inicio dessas regioes e fim o seu valor final
```

```
x=x(:);
x=[0; x; 0];
dx=diff(x);
ini=find(dx == 1);
fim=find(dx == -1)-1;
```

```
function [y,sy]=fmeanrank(x,l,lr)
%FMEANRANK Filtro de media sobre rank
% y=fmeanrank(x,l,lr) Filtragem do sinal x por
% um filtro de media de ordem l, sobre lr
% (l=2*lr+1) pontos ordenados
% l é opcional por defeito é 7.
% Se lr e um vector calcula media sobre lr
% O resultado y é um vector coluna
```

```
if nargin == 1
    l=7;lr=fix(l/2);
elseif nargin == 2
    if rem(l,2) == 0
        l=l+1;
    end
    lr=fix(l/2);
end
```

```
if (length(lr) == 1)
% Calcula dominio para calculo da media
    indice=fix((l+1)/2)+(-fix(lr/2):fix(lr/2));
```

```
elseif (max(lr) > l | min(lr) < 1)
    disp(' FMeanRank: Indice nao valido ')
    return
else
    indice=lr;
end
x=x(:);
lx=length(x);
bl=fix(l/2);
x=[ones(bl,1)*x(1) ; x ; ones(bl,1)*x(lx)];
ind=0:2*bl;
k=1:lx;
kind=k'*ones(1,l)+ones(lx,1)*ind;
y=x(kind');
y=reshape(y,l,lx);
y=sort(y);
sy=std(y(indice,:));
y=mean(y(indice,:));
y=y(:);
sy=sy(:);
```

8.2 Programas de geração de formas fractais

```
function [x,t]=mpm(H,N,seed)
%MPM Midpoint Displacement Method para gerar fBm.
% [x,t]=mpm(H,N,seed) retorna 2^N+1 pontos num
% vector coluna do fBm caracterizado
% pelo parametro H, iniciando a seed de randn
% Opcionalmente retorna o vector temporal
% correspondente ao intervalo [0 1]
```

```
randn('seed',seed)
sigma=1;

lx=2^N+1;
x=zeros(1,lx);
x(lx)=randn(1,1)*sigma;
%x(lx)=0;
```

```
if nargin == 2
    t=(0:lx-1)/(lx-1);
end

for k=1:N;
    delta=sigma*sqrt(1-4^(H-1))*0.5^(H*k);
    dist=2^(N-k);
    poi=((1:2^(k-1))*2-1)*dist+1;
    x(poi)=(x(poi-dist)+x(poi+dist))/2;
    x(poi)=x(poi)+randn(size(poi))*delta;
end
x=x(:);
```

```

function [x,t]=sam(H,N,seed)
%SAM Successive Random Addition Method para gerar
fBm.
% [x,t]=sam(H,N,seed) retorna 2^N+1 pontos num
% vector coluna do fBm caracterizado
% pelo parametro H, iniciando a seed de randn
% Opcionalmente retorna o vector temporal
% correspondente ao intervalo [0 1]

randn('seed',seed)
sigma=1;

lx=2^N+1;
x=zeros(1,lx);
x(lx)=randn(1,1)*sigma;
sigma=sigma*sqrt((1-4^(H-1))/2);

if nargin == 2
    t=(0:lx-1)/(lx-1);
end

for k=1:N;
    delta=sigma*0.5^(H*k);
    dist=2^(N-k);
    poi=((1:2^(k-1))*2-1)*dist+1;
    allpoi=(0:2^k)*dist+1;
    x(poi)=(x(poi-dist)+x(poi+dist))/2;
    x(allpoi)=x(allpoi)+randn(size(allpoi))*delta;
end
x=x(:)-x(1);

```

```

function [x,t]=ssm1(H,N,seed)
%SSM1 Geração de movimento fractional browniano
% por transformação inversa de Fourier
% [x,t]=ssm1(H,N,seed) retorna 2^N+1 pontos num
% vector coluna do fBm caracterizado
% pelo parametro H, iniciando a seed dos geradores
% de numeros aleatorios
% Opcionalmente retorna o vector temporal
% correspondente ao intervalo [0 1]

beta=2*H+1;

N=2^N*2;
randn('seed',seed)
rand('seed',seed)

if nargin == 2
    t=(0:N/2)/(N/2);
end

f=1:N;
rho=f.^(-beta/2); %.*abs(randn(size(f)));
phase=2*pi*(rand(1,N-1)-0.5);

rho=[0 rho rho(N-1:-1:1)];
% normalizacao - teorema de Parseval
% para variancia unitaria
rho=rho*length(rho)/sqrt(sum(rho.*rho));
phase=[0 phase 0 -phase(N-1:-1:1)];

Y=rho.*exp(i*phase);

x=real(iff(Y));

x=x(1:N/2+1);
x=x(:)-x(1);
%plot(0:N,x);

```

```

function [x,t]=ssm2(H,N,fc,seed)
%SSM2 Geração de movimento fractional browniano
% por transformação inversa de Fourier
% [x,t]=ssm2(H,N,fc,seed) retorna 2^N+1 pontos
% num vector coluna do fBm caracterizado
% pelo parametro H, iniciando a seed dos geradores
% de numeros aleatorios.
% fc frequencia (0 < fc < 1)
% de transição de expoente
% Opcionalmente retorna o vector temporal
% correspondente ao intervalo [0 1]

beta=2*H+1;

N=2^N*2;
Nc=fix(fc*N);
randn('seed',seed)
rand('seed',seed)

if nargin == 2
    t=(0:N/2)/(N/2);
end

% Primeiro regime f < fc
f=1:Nc;
rho=f.^(-beta(1)/2); %.*abs(randn(size(f)));
% Segundo regime f > fc
f=(Nc+1):N;
rho=[rho f.^(-beta(2)/2)*Nc^(H(2)-H(1))];

phase=2*pi*(rand(1,N-1)-0.5);

rho=[0 rho rho(N-1:-1:1)];
% normalizacao - teorema de Parseval
% para variancia unitaria
rho=rho*length(rho)/sqrt(sum(rho.*rho));
phase=[0 phase 0 -phase(N-1:-1:1)];

Y=rho.*exp(i*phase);

x=real(iff(Y));
x=x(1:N/2+1);
x=x(:)-x(1);

```

```

function [x,t]=sam2(h,N,nt,seed)
%SAM2 Successive Random Addition Method
% para gerar fBm.
% [x,t]=sam2(H,N,nt,seed) retorna 2^N+1 pontos
% num vector coluna do fBm caracterizado
% pelo parametro H, iniciando a seed de randn.
% nt resolução para transição de regime
% Opcionalmente retorna o vector temporal
% correspondente ao intervalo [0 1]

randn('seed',seed)
randn(1000,1);
sigma=1;

lx=2^N+1;
x=100000*ones(1,lx); % representa nan's
x(1)=0;
x(lx)=randn(1,1)*sigma;

if nargin == 2
    t=(0:lx-1)/(lx-1);
end

% Regime inicial (baixa resoluções
temporais)
delta=sigma*sqrt((1-4^(h(1)-
1))/2)*0.5.^(h(1)*(1:(nt-1)));
% Regime final (altas resoluções temporais)
delta=[delta delta(nt-1)*0.5.^(h(2)*(1:(N-nt+1)))];

for k=1:N
    dist=2^(N-k);
    poi=((1:2^(k-1))*2-1)*dist+1;
    allpoi=(0:2^k)*dist+1;
    x(poi)=(x(poi-dist)+x(poi+dist))/2;
    x(allpoi)=x(allpoi)+randn(size(allpoi))*delta(k);
end

x=x(:)-x(1);

```

8.3 Programas de determinação de dimensões fractais

```
#ifndef MYTYPES_INC
#define MYTYPES_INC

typedef unsigned long int Index;
typedef unsigned char Boolean;
```

```
// Definicao de vectorUL
typedef unsigned int DatUL;
typedef struct vectorul
```

```
{
    DatUL *data;
    Index size;
} VectorUL;
```

```
// Funcoes para VectorUL
VectorUL* newVectorUL(Index);
int delVectorUL(VectorUL*);
```

```
// Definicao de vector
typedef double DatVc;
typedef struct vector
{
    DatVc *data;
    Index size;
} Vector;

// Funcoes para VectorUL
Vector* newVector(Index);
Vector** newVector(Index, Index);
int delVector(Vector*);

DatVc max(Vector*);
DatVc min(Vector*);
DatVc mean(Vector*, Index, Index);
DatVc mean(Vector*);
DatVc std(Vector*);
Vector* diff(Vector*, Index);

#endif
```

```
#ifndef MYTREE_INC
#define MYTREE_INC
```

```
#include "mytypes.h"
```

```
typedef struct treenode
```

```
{
    Index value;
    int level;
    Index count;
// nodo no mesmo nivel
    struct treenode *right;
// nodo no nivel inferior
    struct treenode *down;
} TreeNode;
```

```
// criacao de um objecto do tipo TreeNode
TreeNode* newTree(void);
// inicializacao dos campos nivel e count
int initTree(TreeNode*, int, Index);
// destruicao de um objecto do tipo TreeNode e
// seus derivados
int delTree(TreeNode*);
```

```
// Atribuicao de uma ponto a TreeNode
int addPointTree(TreeNode*, VectorUL*, int);
// Procura de um TreeNode para atribuicao de um
index
int findTree(TreeNode**, Index, int);

// Calculo de nodos no nivel i
float boxCount(TreeNode*, int);
// Calculo da informação de Shannon
float information(TreeNode*, int);
// Subrotina utilizada por information
float infoSub(TreeNode*, int);
// Calculo da correlacao no nivel i
float correlation(TreeNode*, int);
// Subrotina utilizada por correlation
float corrSub(TreeNode*, int);
// Calculo do index baseado no vector
Index calcIndex(VectorUL*);

#endif
```

 Function: addPointTree(tree,point,level)
 Adiciona o ponto point a TreeNode tree ate ao
 seu nivel level
 Autor: Carlos Felgueiras
 *****/

```
int addPointTree(TreeNode *tn, VectorUL *point, int deepTree)
{
    TreeNode *tActual;
    Index index;
    int niv;
```

```
// Actualiza numero de pontos
tn->count++;

// Inicia procura no nivel 1
tActual=tn;
for (nivel=1; nivel <= deepTree; niv++)
{
    index=calcIndex(point);
    findTree(&tActual,index,niv);
}
return 1;
}
```

 Function: findTree(tree,index,nivel)
 Procura o nodo de tree com value index no
 level nivel. Se existir actualiza os seus
 campos. Se nao existir criar um novo nodo
 de tree. O valor de tree vem alterado para
 o novo nodo.
 Autor: Carlos Felgueiras
 *****/

```
int findTree(TreeNode **tn, Index index, int nivel)
{
    TreeNode *aux, *aux2;
    Boolean done;

    // Nao existe nodo no nivel inferior
    if ( (aux=(*tn)->down) == NULL)
    {
        aux=newTree();
        initTree(aux,nivel,index);
        (*tn)->down=aux;
    }
    else
    {
        done=0;
```

```
while ( aux != NULL && !done) {
    // Se aux == NULL, aux2 guarda nodo anterior
    aux2=aux;
    // Nodo ja existente
    if (aux->value == index) {
        // Actualizo numero de pontos
        aux->count ++;
        done=1;
    }
    // Procuro outros nodos no mesmo nivel
    else
        aux=aux->right;
}
// Nao encontrei nodos com value index
if (!done) {
    aux=newTree();
    initTree(aux,nivel,index);
    // Acrescento novo nodo a direita
    aux2->right=aux;
}
*tn=aux;    // Actualiza nodo actual
return 1;
}
```

```
DatUL UL_HALF= (1 << (sizeof(DatUL)*8)-1);
```

 Function: calcIndex(vector)
 Calcula index para o nodo de tree, baseado
 nos valores dos elementos de vector.
 Altera o valor dos elementos do vector.
 Autor: Carlos Felgueiras
 *****/

```
Index calcIndex(VectorUL *point)
{
    Index index, i;
```

```
for (i=0, index=0; i<point->size; i++)
{
    // index *=2;
    index= index << 1;
    // index+= MSB
    index+= (point->data[i] >= UL_HALF);
    // Actualiza point
    point->data[i]= point->data[i] << 1;
}

// Sequencia dos MSB dos elementos do vector
return index;
}
```

```

/*****

```

Function: boxCount(arvore, nivel)

Calculo do numero de hipercubos de dimensao linear $2^{\text{-(nivel)}}$ que contem uma porcao do atrator.

Autor: Carlos Felgueiras

```

*****/

```

```

float boxCount(TreeNode *tn, int nivel)
{

```

```

    if ( tn != NULL && tn->level < nivel)
        return(boxCount(tn->down,nivel) +
                boxCount(tn->right,nivel));
    else {
        if ( tn != NULL && tn->level == nivel)
            return(1.0f+boxCount(tn->right,nivel));
        else
            return(0.0f);
    }
}

```

```

/*****

```

Function: information(arvore,nivel)

Calculo da informacao de Shannon, baseada nas frequencias relativas da ocorrencia de pontos do atrator num hipercubo de dimensao linear $2^{\text{-(nivel)}}$.

Autor: Carlos Felgueiras

```

*****/

```

```

float information(TreeNode *tn, int nivel)
{
    // tn é o nodo inicial da arvore (tn->level = 0)
    // tn->count conta o numero de pontos (N)
    // H=log(N)-sum(ni*log(ni))/N= -sum(pi*log(pi)),
    // pi=ni/N, em bits H=H/log(2)

    return(float)((log((double)tn->count) -
        infoSub(tn->down,nivel)/tn->count) /
        log((double)2.0));
}

```

```

/*****

```

Function: infoSub(arvore,nivel)

Subrotina utilizada por information

Autor: Carlos Felgueiras

```

*****/

```

```

float infoSub(TreeNode *tn, int nivel)
{
    if ( tn != NULL && tn->level < nivel)
        return(infoSub(tn->down,nivel) +
                infoSub(tn->right,nivel));
    else
    {
        if ( tn != NULL && tn->level == nivel)
            return((float)(tn->count *
                log((double)tn->count)) +
                infoSub(tn->right,nivel));
        else
            return(0.0f);
    }
}

```

```

/*****

```

Function: correlation(arvore,nivel)

Calculo do "comprimento de correlacao" entre os pontos do atrator, baseado nas frequencias relativas de ocorrencia dos mesmos num hipercubo de dimensao linear $2^{\text{-(nivel)}}$.

Autor: Carlos Felgueiras

```

*****/

```

```

float correlation(TreeNode *tn, int nivel)
{
    // tn é o nodo inicial da arvore (tn->level = 0)
    // tn->count conta o numero de pontos (N)
    // C= sum(ni^2)/N^2= sum(pi^2), pi=ni/N

    return(corrSub(tn->down,nivel) /
        ((float)tn->count*tn->count));
}

```

```

/*****

```

Function: corrSub(arvore,nivel)

Subrotina utilizada por correlation

Autor: Carlos Felgueiras

```

*****/

```

```

float corrSub(TreeNode *tn, int nivel)
{
    if ( tn != NULL && tn->level < nivel)
        return(corrSub(tn->down,nivel) +
                corrSub(tn->right,nivel));
    else
    {
        if ( tn != NULL && tn->level == nivel)
            return((float)(tn->count * tn->count +
                corrSub(tn->right,nivel));
        else
            return(0.0f);
    }
}

```

```

function [L,k]=higushi(x)
% HIGUSHI      Determinação da dimensão fractal
%              pelo metodo de Higushi
% [L,k]=higushi(x) Retorna os comprimentos L de
%              versoes subamostradas k vezes do sinal fractal x.
% L é uma matriz 2*n sendo L(1,:) a media dos
% comprimentos e L(2,:) a sua variância.
% A dimensão fractal é determinada a partir da
% relação  $L(1,:)=A*k^{(-D)}$ 

x=x(:);          % x é agora um vector coluna
lx=length(x);
kmax=lx/3;
%      numero maximo de bins = 20!!!!
jmax=min(fix(log2(kmax))*4+1,27);
j=1:jmax;
k=[1 2 3 4 fix(2.^(j-1)/4)];

for j=1:length(k)
%      Nao sao aproveitados alguns pontos da
%      cauda do sinal. Higushi aproveita-os
ind=1:k(j):(fix(lx/k(j))*k(j));
for k2=1:k(j)
%      Sinais subamostrados k vezes
y(:,k2)=x(ind+k2-1);
end
%      Distâncias entre dois pontos sucessivos
%      dos sinais
y=abs(diff(y));
%      Comprimentos dos sinais
y=sum(y);
%      Correção para os pontos não utilizados
y=y*length(x)/(k(j)*length(ind));
L(1,j)=mean(y); L(2,j)=std(y);
y=[];
end

```

```

function [S, f]=cglsa(x,nfft)
%CGSA Coarse graining spectral analysis
% [S, f]=cglsa(x,nfft)

if nargin < 2
    nfft=256;
end
% Transformada com potencia de dois pontos
nfft=2^fix(log2(nfft));

x=x(:);
ntotal=length(x);
% Sobreposicao a +/- 50%
lag=nfft/2-1;
noverlap=nfft-lag;

% Numero util de pontos
nutil=ntotal-3*nfft;
nsubset=fix((nutil-noverlap)/(nfft-noverlap));

isec=0:nfft-1;
ifft=1:nfft/2+1;
win=hanning(nfft);
KMU=norm(win)^2;

% Calculo do primeiro vector de incrementos da fase
% do espectro cruzado a partir das duas primeiras
% seccoes do sinal
xs=zeros(nfft,6);
y=zeros(nfft,6);
Sc=zeros(length(ifft),4);
for k=0:1
    xs(:,1+3*k)=detrend(x(k*lag+2*isec+1),0).*win;
    xs(:,2+3*k)=detrend(x(k*lag+isec+1),0).*win;
    xs(:,3+3*k)=detrend(x(k*lag+4*isec+1),0).*win;
end
y=fft(xs);

Sc(:,1:2)=y(ifft,[1 4]).*conj(y(ifft,[2 5]))/KMU;
Sc(:,3:4)=y(ifft,[1 4]).*conj(y(ifft,[3 6]))/KMU;
theta=angle(Sc(:,[1 3 2 4]));

xs(:,4:6)=[];
y(:,4:6)=[];
S=zeros(length(ifft),3);

% Calculo dos espectros a partir das restantes seccoes
for k=2:nsubset-1
    xs(:,1)=detrend(x(k*lag+2*isec+1),0).*win;
    xs(:,2)=detrend(x(k*lag+isec+1),0).*win;
    xs(:,3)=detrend(x(k*lag+4*isec+1),0).*win;

    y=fft(xs);
    Sc(:,[1 3])=Sc(:,[2 4]);
    Sc(:,2)=y(ifft,1).*conj(y(ifft,2))/KMU;
    Sc(:,4)=y(ifft,1).*conj(y(ifft,3))/KMU;

    delta=theta(:,3:4)-theta(:,1:2);
    theta(:,1:2)=theta(:,3:4);
    theta(:,3:4)=angle(Sc(:,[2 4]));

    Slin=(Sc(:,[2 4]).*exp(i*(pi/2-delta)));
    Slin=abs(real(conj(Slin).*Sc(:,[1 3]))) /
        (abs(Sc(:,[1 3]))+eps); % |<A,B>|

    S(:,1)=S(:,1)+y(ifft,1).*conj(y(ifft,1))/KMU;
    S(:,[2 3])=S(:,[2 3])+Slin;
end

S(:,2:3)=S(:,2:3)*2;
S=S/(nsubset-2);
S(:,4)=sqrt(S(:,2).*S(:,3));

f=(ifft-1)/nfft*2; % Fsampling = 2

```

```
#define SQRT1_2 0.707106781
```

```
int findExp(double val)
```

```
{
    int expo;
    double mant;
```

```

    mant= frexp(val, &expo);
    expo= 2*expo - (mant < SQRT1_2);
    if (expo < 0) expo=0;

    return expo;
}

```

```

extern Index histo[DIMMAX+1][NBINS+1];

/*****
Function: functionPG(dat, delay, dmin, dmax)
Calcula o histograma das distância entre
pontos do atractor construido a partir do
vector dat e delay, para todas as dimensoes
do espaço de fase entre dmin edmax.
Autor: Carlos Felgueiras
*****/

void functionPG(Vector *dat, Index delay,
                Index dMin, Index dMax)
{
    Index k1, k2, kFixP, kVarP, dim, nDataUtil,
          nPoiOri, nP;
    double distMax, dist;
    int expoente, incr;
    DatVc mx, Mx, *x;

    for (k1=0; k1 <= DIMMAX; k1++)
        for (k2=0; k2 <= NBINS; k2++)
            histo[k1][k2]=0;
    for (k2=0; k2 <= NBINS; k2++)
        histo[dMin-1][k2]=k2;

    x=dat->data;

    // Numero de pontos do atractor
    nDataUtil= dat->size-delay*(dMax-1);
    // Numero de pontos como origem.
    nPoiOri=NPOINTMAX/nDataUtil;

    for (nP =0; nP < nPoiOri; nP++) {
        // Escolha aleatória das origens
        k1=(Index)((float)rand()*(nDataUtil-1)/RAND_MAX);

        for (k2 =0; k2 < nDataUtil; k2++)

            srand((unsigned)time( NULL ) );
            {
                kFixP=k1;
                kVarP=k2;
                dist=0.0;
                for (dim=1; dim < dMin; dim++) {
                    dist+= (x[kFixP]-x[kVarP])*(x[kFixP]-x[kVarP]);
                    kFixP+= delay;
                    kVarP+= delay;
                }
                for ( ; dim <= dMax; dim++) {
                    dist+= (x[kFixP]-x[kVarP])*(x[kFixP]-x[kVarP]);
                    kFixP+= delay;
                    kVarP+= delay;
                    expoente=findExp(dist);
                    histo[dim][expoente]++;
                }
            }
        }

    // Calcula histograma acumulado
    for (k1=dMin; k1 <= dMax; k1++)
        for (k2=1; k2 <= NBINS; k2++)
            histo[k1][k2]+= histo[k1][k2-1];
}

```

8.4 Publicações

O artigo das páginas seguintes foi apresentado no IV Congresso Português de Engenharia Biomédica realizado em Aveiro, 18 e 19 de Abril de 1996 e consta das respectivas Actas.

On the Reconstruction of Fetal Heart Rate Sequences

C. S. Felgueiras^(*), J. P. Marques de Sá^(*), J. Bernardes^(**), S. Gama^(*)

^(*) Faculdade de Engenharia da Universidade do Porto

^(**) Faculdade de Medicina da Universidade do Porto

Abstract - The fetal heart rate (FHR) sequence is a component signal of the cardiotocographic signals set, obtained by fetal monitors and used as a perinatal diagnostic tool¹ in obstetric care services.

Among non-invasive techniques for measuring the FHR is the application of the autocorrelation technique to Doppler ultrasounds². However this technique originates, frequently, some artefacts.

In the present work we present a method for detecting and correcting those artefacts, based only on the FHR sequence supplied by the fetal monitor.

I. INTRODUCTION

We assume, as usual, that the FHR sequence is the only data available and based on it we interpret the different kinds of simple artefacts (double counting, half counting artefacts and spikes) which may be present, resulting from insufficiencies of the autocorrelation technique of the fetal monitor, as consequence for instance of fetal movements.

Since the FHR sequence is the inverse of the temporal displacement between two consecutive beats, one can always determine the time at which each beat occurred. This is illustrated at figure 1, where an arrow represents one beat and the square dots show the respective FHR sequence (the discrete signal provided by any fetal monitor).

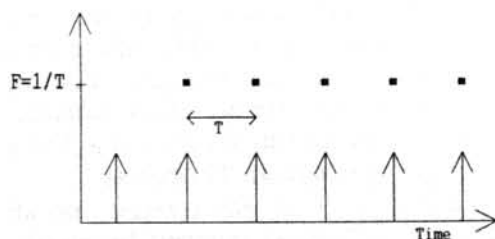


Figure 1 Constant FHR Sequence

In the following we will consider that the original FHR sequence is constant. Under this hypothesis, a half counting (missing one beat) will be detected as a valley in the measured sequence. The occurrence of this type of artefact is illustrated in figure 2, where the dotted arrow marks the beat we have missed and

the open square dot a sample not present in the measured sequence.



Figure 2 Half Counting

In figure 3, we show the situation corresponding to double counting. In this case, an extra beat (artefact), represented by the dashed arrow, is detected. The consequence of this extra beat is the appearance in the measured sequence of a peak with an extra sample showed by the grey filled square dots in the figure.

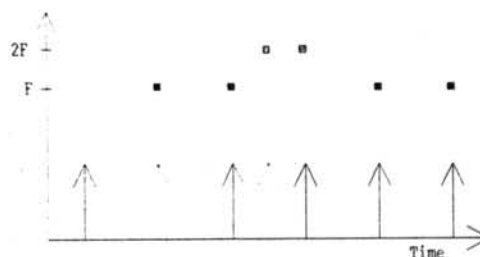


Figure 3 Double Counting

A spike occurs when we miss the detection of a beat, illustrated in figure 4 as a dotted arrow, but we detect an artefact in the neighbourhood showed as a dashed arrow. So, in the measured sequence, we have a peak followed by a valley, or vice-versa, with the same number of samples as in the original sequence, with the values F_1 and F_2 satisfying:

$$\frac{1}{F_1} + \frac{1}{F_2} = \frac{2}{F}$$



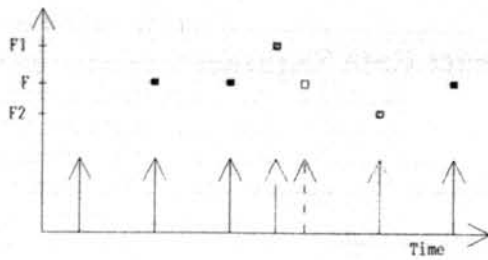


Figure 4 Spikes

For the three cases we have described, the correction of the artefacts is constrained by the condition:

The time axis must not be deformed.

Therefore, the condition that the values of the artefacts samples (grey filled squares in the above figures) $f(i):i=1,...,N$ and the F value of the original sequence must satisfy is:

$$\sum_{i=1}^N \frac{1}{f(i)} = \frac{M}{F}, \text{ for some integer } M.$$

Or rewriting in order to find M :

$$1. \quad M = \sum_{i=1}^N \frac{F}{f(i)}.$$

So, we must substitute the N samples of the artefact by M samples with value F .

Since in real sequences we cannot always find a constant sequence to substitute for the artefact and, most importantly, it is possible to have a mixing of all simple artefacts to form compound artefacts, another approach must be considered:

2. Because the half counting is the only case where a single sample "contains" all the information in presence of a good estimate of F , it is also the only case about which we can be sure of what happened.

Therefore we first consider the reconstruction of this type of artefacts.

II. METHOD

The algorithm we propose is divided in two components: artefacts detection and artefacts correction.

A. Artefacts detection

The first idea that comes to mind for detecting artefacts consists in the application of non-linear filters (median filters and morphological filters), to

the measured sequence, so that we can detect the artifacts by the difference between the original and filtered sequence. We tried this approach without much success given the difficulty of having a filter coping with all possible artefacts situations and obtaining usefull information from the detected differences.

Another approach is to study the increments of the measured sequence, as is done in SisPorto¹, which performs a certain degree of artefact detection and reconstruction.

The approach we have developed takes into consideration the appearance of sequence discontinuities and is as follows:

Let $f(i)$, $i=1,...,N$ be the measured sequence. The increment to any given sample j is given by:

$$d(j) = f(j+1) - f(j), j=1,...,N-1.$$

To detect the artefacts we first classify each sample as been valid or not valid according to a discontinuity rule:

- If $|d(j)| > D$ classify the sample j and $j+1$ as not valid.

It can be shown that this rule is equivalent to the following one:

- The sample j is classified as valid if:

$$\forall i \in \{0,1\}: |f(j+i) - f(j+i-1)| \leq D.$$

So, for validation of a sample j we use the values of samples in the window $\{j-1, j, j+1\}$.

This rule is different of that used by SisPorto in two aspects:

- The width of the window used to validate samples was five in SisPorto (differences with the four neighbours). We use a width of three, because it is the minimum width necessary for delimitation of simple artefacts. Larger widths will delimit larger compound artefacts, but there will be, allways, artefacts which are not totally delimited. So it is better to use a width of three and to find a good classification scheme for the artefacts.
- The parameter D is not constant for all sequences as in SisPorto.(10 beat/min). Because the different classes of FHR sequences admit different distributions for the increments $d(j)$ it is convenient to choose a specific value for D for each sequence. The determination of this value is based on the histogram of the absolute values of $d(j)$. D is the first local minimum of the histogram such that the

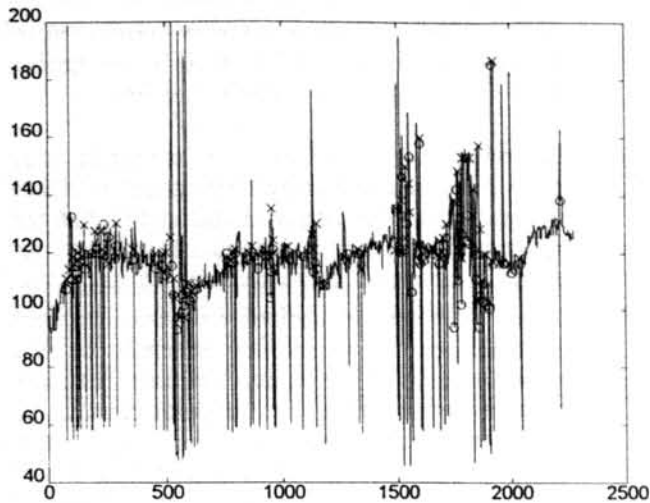


Figure 5 Artifacts on a measured sequence.

"x" marks the beginning and "o" the end of an artifact.

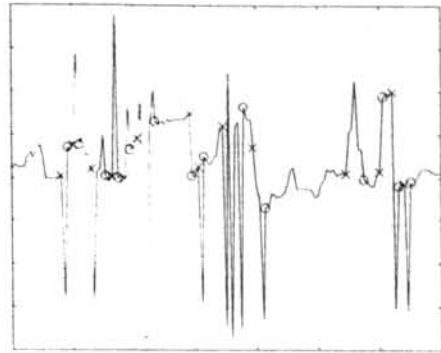


Figure 5a Zoom of figure 5 on region 1480:1620

number of occurrences of $\{D-1, D, D+1\}$ is inferior to 1% of the total number of samples.

We then identify the artefacts by finding the beginning and end of subsequences of non valid samples. The results of this procedure are illustrated in figure 5, for a particular signal showing examples of the different types of artefacts.

Next, we proceed with the correction of the artefacts if necessary.

B. Artefacts correction

Let $b(j)$ and $e(j)$, $j=1, \dots, N_s$ be the beginning and end of the N_s artefacts returned by the previous processing.

Now we establish the following two classes of artefacts:

- C1 Artefacts that satisfy:

$$|f(b(j)) - f(e(j))| \leq D, j=1, \dots, N_s.$$

- C2 Other artefacts.

On the correction algorithm, we first correct the artefacts of class C1, since there is a high degree of continuity between their limits. Taking this into account, we take the mean of the values at their limits as an estimate of the real value of the interior points of the artefact. Next we find the number of samples missing or in excess by taking the quotient between this estimated sequence and the measured sequence, as shown in figure 6. Following the approach mentioned in point 2 of the Introduction, we first find the values of the quotient which are near two (half counting) or three (found in some sequences). For these samples we replace the mean

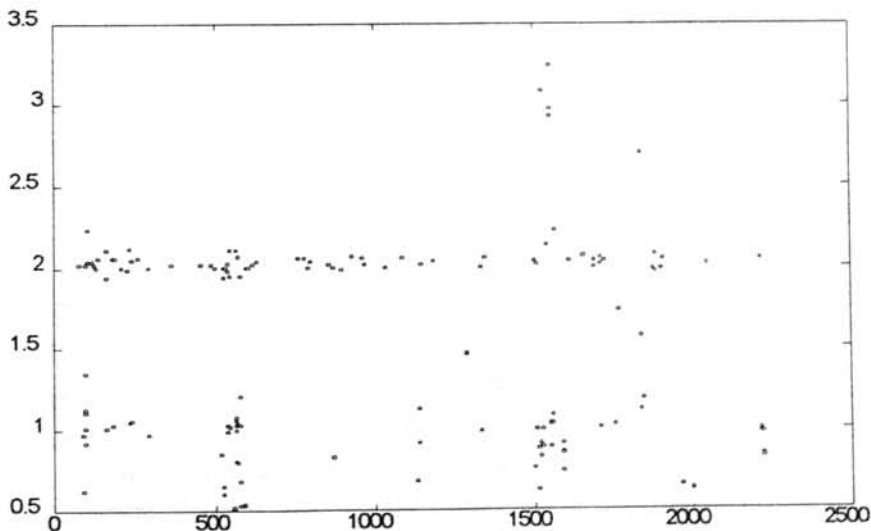


Figure 6 Quocient for artifacts of type C1

value of the estimated sequence by two or three samples with the double or triple of the value of the measured sequence, respectively.

For the other samples of the artefact we maintain the estimated value, and find the real number of samples, which must fill the artefact interval, by finding the nearest integer to the cumulative sum of contiguous subsequences of non integer quotients, according to the formula of point 1 of the Introduction.

For artefacts of class C2 we introduce four subclasses based on the rules:

- C21 Artefacts which satisfy:

$$|f(b(j)) - f(e(j))| \leq 2D.$$

- C22 Pairs of artefacts not included in C21 which satisfy:

$$\begin{cases} |f(e(j)) - f(b(j-1))| \leq D \\ |f(b(j)) - f(e(j-1))| \leq D \end{cases}$$

- C23 Artefacts not included in C21 or C22 such that the linear interpolation between $(b(j), f(b(j)))$ and $(e(j), f(e(j))) - F(b(j): e(j))$, satisfy

$$\sum_{i=b(j)}^{e(j)} |f(i) - F(i)| \leq D.$$

- C24 Other artifacts.

This classification scheme must be carried out in the sequence C21, C22, C23 and C24 and applied only to the artifacts not previously classified.

For classes C22 and C23 to be consistent it is necessary that after an artefact with a fast rise there is another with a fast decay to the initial value, or vice-versa, otherwise those artefacts will be classified as C23, and will be considered real features of the sequence and not corrected. This kind of hypothesis is not always verified, so the discriminatory rules for classes C22 and C23 must be optimised. This type of misclassification is showed in figure 7. Artefact samples are represented by filled circles and this classification scheme works well with the first and second artefacts, but classifies the other artefacts as C23 and leave them uncorrected.

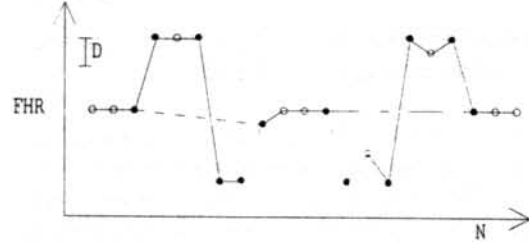


Figure 7 Errors of classification of C22 and C23

Correction is done only for classes C21, C24 and C22, after fusion of artefacts of this class, and the procedure is the same one as on the second run for correction of artefacts of class C1, with the estimated value being the mean of the values of the limits of the artefacts. Because of larger gaps between those

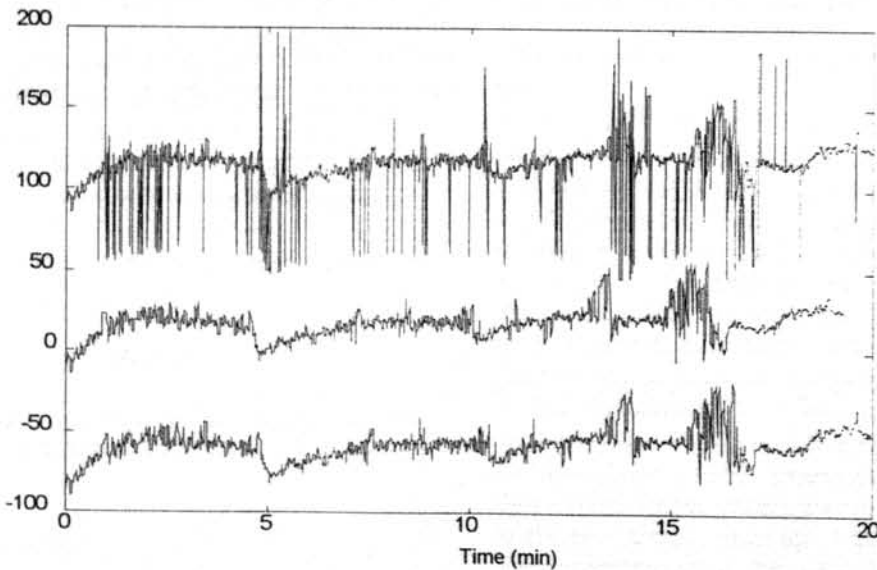


Figure 8 Results of reconstruction, from top to bottom is the original sequence, the reconstructed sequence without correction of the number of samples and the reconstructed sequence (the last two were shifted downwards by 100 and 175, respectively).

values we then transform the flat region between those limits by a line joining them.

III. CONCLUSION

The result of reconstruction of FHR sequences by the method we have just presented is shown in figure 8, for a particular sequence. It is evident that the method is capable of detecting the artefacts present in the sequence. And that the condition underneath the method proposed is fully accomplished, since temporal duration of the reconstructed sequence is only 0.05% less than the actual temporal size of the original sequence. For comparison, we present a corrected sequence, by interpolation of the artefacts without correction of the number of samples. For this sequence the temporal duration is 3.5% less the original sequence, which corresponds to 3/4 of a minute.

We tried the method in other sequences with quite satisfactory results. Some of the remaining found errors were related to the problem, previously mentioned, about C22 and C23, thus demanding an improvement of the classification scheme, which is under current progress.

ACKNOWLEDGEMENTS

C.S.F. is supported by PRAXIS XXI/ BM/ 2340/ 94.

REFERENCES

- [1] Marques de Sá JP, Moura AC, Bernardes J and Pereira Leite L, "Análise Automática de Sinais Cardiotocográficos", *Acta Médica Portuguesa* 5: 28, 1992
- [2] Tuck DL, "Improvement in Doppler Ultrasound Human Foetal Heart Rate Records by Signal Correlation", *Medical & Biological Engineering & Computing* 20: 357, 1992
- [3] Dawes GS, Moulden M and Redman CWG, "Limitations of Antenatal Fetal Heart Rate Monitors", *Am. J. Obstet. et Gynecol.* 162: 170, 1990