Clarice Favaretto Salvador

O Problema da Recuperação da Fase da Transformada de Fourier: Novos Resultados

Prof. Dr. Alvaro Rodolfo De Pierro Orientador

> Prof. Dr. Nir Cohen Co-orientador

Tese apresentada ao Instituto de Matemática Estatística e Computação Científica, UNI-CAMP para obtenção do título de Doutoramento em Matemática Aplicada.





O Problema da Recuperação da Fase da Transformada de Fourier: Novos Resultados

Este exemplar corresponde a redação final da tese devidamente corrigida e defendida pela Sra.**Clarice Favaretto Salvador** e aprovada pela Comissão Julgadora.

Campinas, 10 de Dezembro de 1997.

Prof. Dr. _____

Orientador

Niv Cohi Prof. Dr. ____ Co-orientador

Tese apresentada ao Instituto de Matemática Estatística e Computação Científica, UNICAMP como requisito parcial para a obtenção do Título de DOUTOR em MA-TEMÁTICA APLICADA.

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO IMECC DA UNICAMP

Salvador, Clarice Favaretto
O problema da recuperação da fase da transformada de Fourier: novos resultados / Clarice Favaretto Salvador -- Campinas, [S.P. :s.n.], 1997.
Orientador : Alvaro Rodolfo De Pierro, Nir Cohen Tese (doutorado) - Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.
1. Espectroscopia por transformada de Fourier. 2. Métodos iterativos (Matemática). 3. Astrofísica. I. De Pierro, Alvaro Rodolfo. II. Cohen, Nir. III. Universidade Estadual de Campinas. Instituto de Matemática. Estatística e Computação Científica. IV. Título.

Sa38p

Tese de Doutorado defendida e aprovada em 10 de dezembro de 1997

Pela Banca Examinadora composta pelos Profs. Drs.

Prof (a). Dr (a). ÁLVARO RODOLFO DE PIERRO

Prof (a). Dr (a). SÉRGIO SHIGUEMI FURUIE

Prof (a). Dr (a). JÖRG DIETRICH WILHELM SCHLEICHER

03.11

Prof (a). Dr (a). OSWALDO BAFFA FILHO

Mian Jun - Junits Prof (a). Dr (a). LÚCIO TUNES DOS SANTOS

A meu filho João Pedro e a meu marido Nivaldo.

.

AGRADECIMENTOS

Ao professor Alvaro Rodolfo de Pierro, pela proposta do presente trabalho e pela sua orientação na elaboração do mesmo.

Ao professor Nir Cohen, pelo interesse em ajudar neste trabalho e por suas maravilhosas sugestões.

Ao professor Petrôneo Pulino, por conceder a subrotina que calcula a transformada rápida de Fourier.

Aos professores, funcionários e colegas do Departamento de Matemática Aplicada, pelo apoio e incentivo.

A meu marido Nivaldo, por sua paciência durante o tempo de elaboração deste trabalho.

A todos os amigos, pelo incentivo.

A FAPESP e a CAPES pelo apoio financeiro.

Conteúdo

1	INT	ROD	UÇÃO	4
	1.1	A Uni	cidade	6
		1.1.1	Recuperação da Fase	6
		1.1.2	Deconvolução Cega	9
	1.2	Aplica	ções	10
		1.2.1	Astrofísica, Interferometria Especular	10
		1.2.2	Cristalografia	11
	1.3	Panor	ama desta Tese	15
2	MÉ	TODC	S	17
	2.1	Métod	los para Tipos Especiais de Objetos	17
		2.1.1	Holografia	17
		2.1.2	Difference Fourier Syntheses	18
		2.1.3	Produto de Autocorrelações	18
	2.2	Métod	los para Objetos Gerais	19
		2.2.1	Newton-Raphson	20
		2.2.2	Gradientes	21
		2.2.3	Método do teorema da amostragem	21
		2.2.4	Simulated Annealing	22
		2.2.5	Levenberg-Marquardt	23

	2.3	Métodos Iterativos	23
		2.3.1 Métodos iterativos para recuperar a fase da transformada de Fourier	23
		2.3.2 Métodos Iterativos para Recuperar a Amplitudede da Transformada	
		de Fourier	29
		2.3.3 Métodos Iterativos para Deconvolução de Funções	30
		2.3.4 Tipos de Estagnações e Métodos para Sobrepô-las	3 1
3	NO	VOS MÉTODOS	34
•	3 .1	Função de Relaxação	34
	3.2	Projeções Generalizadas Modificado	35
4	\mathbf{CA}	RACTERIZAÇÃO DO CONJUNTO DOS OBJETOS COM AM-	
	\mathbf{PL}	TUDE DA TRANSFORMADA DE FOURIER DADA.	38
	4.1	Sinais Unidimensionais	39
	4.2	Imagens	42
5	\mathbf{PR}	OPRIEDADES DE CONVERGÊNCIA	50
	5.1	Limitação dos Algoritmos	51
		5.1.1 Error-Reduction	51
		5.1.2 Output-Output	51
		5.1.3 Input-Output Híbrido	51
		5.1.4 Input-Output	72
	5.2	Pontos Fixos	72
		5.2.1 Error-Reduction	73
		5.2.2 Output-Output	73
		5.2.3 Input-Output	73
		5.2.4 Input-Output Híbrido	74
6	UM	MÉTODO AUTOMÁTICO PARA SUPERAR A ESTAGNAÇÃO	75

	6.1	O Novo Método	77				
7	EX	EXPERIMENTOS					
	7.1	Linguagem e pacotes utilizados	79				
	7.2	Imagens e seus formatos	79				
	7.3	Implementação dos algoritmos	80				
		7.3.1 Dados	80				
		7.3.2 Imagens iniciais	81				
•		7.3.3 Testes para parada e acompanhamento dos programas	81				
	7.4	Algoritmos de Fienup utilizando informação sobre a caracterização	82				
	7.5	Resultados	82				
		7.5.1 Exemplos	84				
8	8 CONCLUSÕES						

Capítulo 1

INTRODUÇÃO

Este trabalho é um estudo de alguns aspectos do problema de recuperar uma função dada informação parcial sobre a mesma e sobre sua transformada de Fourier. Este problema matemático surge em inúmeras e importantes aplicações, por exemplo nas áreas de microscopia eletrônica [32], interferometria especular [4], cristalografia de raio X [44] e astrofísica [35]. Descreveremos alguns exemplos importantes com mais detalhes posteriormente. Não discutiremos o problema relacionado com a existência de solução pois na prática admite-se que existe; alguns resultados relacionados com a estabilidade dessa solução com relação aos dados podem ser encontrados em [45]. Comentaremos a questão da unicidade da solução e trataremos principalmente os métodos existentes para a recuperação de tais funções para alguns problemas específicos.

Para uma função $f : \mathbb{R}^M \to \mathbb{C}$ ou $f : \mathbb{R}^M \to \mathbb{R}$, dependendo do caso, podemos considerar essencialmente os seguintes problemas:

RECUPERAÇÃO DA FASE:

(P1) Recuperar f dadas as amplitudes de f e de sua transformada de Fourier F; (P2) Recuperar f dada a amplitude de sua transformada de Fourier F (|F|) e alguma condição convexa sobre f (ou seja, expressa na forma de um conjunto convexo ao qual f deve pertencer);

RECUPERAÇÃO DA amplitude:

(P3) Recuperar f dada a fase de sua transformada de Fourier F e alguma condição convexa sobre f;

e para funções $f: \mathbb{R}^M \to \mathbb{R}$ e $g: \mathbb{R}^M \to \mathbb{R}$ consideramos o problema de

DECONVOLUÇÃO CEGA:

(P4) Recuperar $f \in g$ dada a convolução $f \star g$ das mesmas.

Seja agora uma função f(x) onde x é um vetor M-dimensional e f é integrável. A transformada de Fourier (TF) de f(x), é dada por

$$F(u) = \int_{\Omega} f(x) \exp(-i2\pi u \cdot x) dx = |F(u)| \exp[i\psi(u)],$$
(1.1)

onde $\psi(u)$ é a fase e Ω é \mathbb{R} , \mathbb{R}^2 ou \mathbb{R}^3 , dependendo da aplicação. Nos concentraremos no caso M=2, que é o caso de imagens, porém existem aplicações onde M=1 (em sinais) ou M=3 (por exemplo em cristalografia). A imagem é dada pela transformada de Fourier inversa (TFI) da visibilidade,

$$\mathcal{F}^{-1}(F)(x) = f(x) = \int_{\Omega} F(u) \exp(i2\pi u \cdot x) du$$
(1.2)

e reconstruir a imagem é imediato quando é possível medir F(u) (ver [7] e [42]). Um problema de fase (P1 ou P2) aparece quando é possível medir somente a amplitude |F(u)|. Como a imagem é dada pela equação (1.2) recuperar a imagem é equivalente a reconstruir a fase $\psi(u)$ de F(u).

Com o auxílio do Teorema da Convolução é possível determinar a autocorrelação de f(x) a partir da amplitude de F(u), isto é, $\mathcal{F}^{-1}(|F(u)|^2) = \mathcal{F}^{-1}(F(u)F^*(u)) = f(x) \star f^*(-x) = \int_{-\infty}^{+\infty} f(\alpha) f(x + \alpha) d\alpha$ onde \star e * denotam, respectivamente, a convolução e a conjugação complexa.

Na prática [7] os problemas são discretizados e, no caso 2-dimensional (M=2), são

utilizadas as transformadas de Fourier discretas (TFD)

$$F(u) = \sum_{x \in \Gamma} f(x) \exp(-i2\pi u \cdot x/N)$$
(1.3)

e sua inversa (TFDI)

$$f(x) = N^{-2} \sum_{u \in \Omega} F(u) \exp(i2\pi u . x/N),$$
(1.4)

onde Ω e Γ representam malhas com N^2 pontos. Ambas são computadas utilizando o método da transformada rápida de Fourier [7].

Segundo a definição de Hadamard [3] um problema é bem posto se a solução

- (i) existe
- (ii) é única

(iii) depende continuamente dos dados.

Se o problema viola qualquer uma destas condições ele é mal posto. Aqui violações da condição (i) não são importantes pois se trata de um fenômeno real cuja solução existe. A não existência de solução para o modelo não implica que a quantidade física não exista e sim se deve a um modelo impreciso ou a dados ruidosos. Ambas as possibilidades são exemplos extremos de dependência descontínua nos dados (violação da condição (ii)). Discutiremos as condições para a unicidade da solução.

1.1 A Unicidade

1.1.1 Recuperação da Fase

É claro que existem algumas propriedades da função que são perdidas quando não é possível medir a fase da TF. Por exemplo, quando conhecida somente a amplitude da TF da função (**problema de recuperação de fase puro**), se f(x) é uma solução do problema também são soluções kf(x) com |k| = 1, $f(x + \alpha)$ e $f^*(-x)$. Como do ponto de vista prático estas são ambiguidades aceitáveis (fator de fase constante, translações ou inversão na origem) a unicidade é discutida a menos dessas funções que chamaremos de associadas triviais.

Se a imagem é uma função complexa arbitrária o problema da fase tem uma infinidade de soluções pois pode-se associar qualquer função de fase a amplitude conhecida. O'Neill e Walter [36] parecem ter sido os primeiros a reconhecer que o problema posto assim não corresponde a qualquer situação física real. Em qualquer caso prático, existem restrições na forma da imagem. Por exemplo, em sistemas óticos a imagem tem suporte finito e é usualmente não negativa.

Akutowics, Hofstetter e Walther (ver [1], [50] e [25]), verificaram que o problema contínuo da recuperação da fase (P1 ou P2) unidimensional (M=1) em geral não tem solução única. Walther mostrou que se o suporte de f(x) é finito este problema tem um número infinito mas enumerável de soluções. Ele observou ainda que outras restrições, como por exemplo a não negatividade de f(x), devem reduzir o número de soluções, porém, não é claro qual o grau da redução.

Para o caso contínuo em duas ou mais dimensões a análise é diferente pois o conjunto das funções inteiras não é um anel de fatoração e a expansão de Hadamard não se extende a funções inteiras multidimensionais (teoria utilizada por Akutowics, Hofstetter e Walther). Sanz [45] deu a seguinte condição, suficiente mas não necessária, para a unicidade de solução:

Se $F : \mathbb{R}^n \to \mathbb{C}$ é uma função de banda limitada (sua TFI, f, tem suporte compacto) e $F(z_1, z_2, ..., z_n)$, sua única extenção inteira do tipo exponencial a \mathbb{C}^n , é irredutível em \mathbb{C}^n , então f pode ser unicamente reconstruída (a menos de associadas triviais) a partir de $|F(u_1, u_2, ..., u_n)|$.

No caso do problema unidimensional discreto, Bruck e Sodin [8] mostraram, utilizando o sistema gerado pela autocorrelação e a Z-transformada de f [40], que se f é de suporte finito, então existem no máximo 2^N soluções (N=número de amostras).

Ainda em 1979, Bruck e Sodin [8] sugeriram a possibilidade de reconstrução única de uma função f de suporte finito no problema discreto multidimensional. Em 1982, Hayes [22] deu uma caracterização completa desse problema incluindo um limitante superior para o número de soluções. Um resultado de extrema importância mostra que para $M \ge 2$ o problema de recuperação de fase tem solução quase sempre única. Mais especificamente, considerando que a Z-transformada de uma seqüência m-dimensional $x(n_1, ..., n_m)$ é definida por

$$X(z_1, ..., z_m) = \sum_{n_1} \sum_{n_2} ... \sum_{n_m} x(n_1, ..., n_m) z_1^{-n_1} ... z_m^{-n_m}$$

e é dita simétrica se para algum vetor k de números inteiros positivos $X(Z) = \pm Z^{-k}X(Z^{-1})$ Hayes demonstrou que:

Se a Z-transformada $A(z_1, z_2, ..., z_n)$ de uma seqüência multidimensional $a(j_1, ..., j_n)$ é irredutível e não simétrica (como um polinômio em n variáveis), então a seqüência "a" é unicamente determinada a partir da amplitude de sua transformada de Fourier, exceto por associadas triviais.

Outro resultado neste sentido foi obtido por Sanz ([45]):

Qualquer seqüência real $a(j_1, ..., j_n)$, $0 \le j_i \le N-1$, i=1,...,n, $n \ge 2$, a menos de uma variedade algébrica em $\mathbb{R}^N \times \mathbb{R}^N \times ... \times \mathbb{R}^N$, pode ser unicamente reconstruída (a menos de associadas triviais) a partir da amplitude de sua série de Fourier.

Observamos que embora esse resultado seja para seqüências reais, ele pode ser extendido para seqüências de números complexos e então se aplica ao problema (P1).

Em 1980, Hayes, Lim e Oppenheim [23] mostraram que uma seqüência unidimensional de comprimento finito (suporte finito) é unicamente reconstruída, a menos de um fator escalar, pela fase ou a tangente da fase de sua transformada de Fourier se sua Z-transformada possue todos os seus zeros no círculo unitário ou em pares recíprocos ($z e z^{-1}$). Hayes [22]) estendeu este resultado para seqüências multidimensionais com uma restrição que envolve a noção de Z-trasformada simétrica. Assim uma seqüência multidimensional de suporte finito é unicamente reconstruída, a menos de um fator escalar, a partir da fase ou da tangente da fase de sua transformada de Fourier se sua Z-transformada não possui fatores simétricos.

1.1.2 Deconvolução Cega

A convolução g(x) de duas funções, $f(x) \in h(x)$, é definida pela integral

$$g(x) = \int_{-\infty}^{\infty} f(x')h(x - x')dx'.$$
 (1.5)

O problema de deconvolução cega consiste em reconstruir a função f(x) a partir de g(x)sem o conhecimento de h(x). Este problema é altamente indeterminado (com nenhuma unicidade de solução) a menos que se tenham mais informações sobre as funções envolvidas. Se, por exemplo, uma das funções é Hermitiana é possível reconstruir analiticamente a fase da transformadada de Fourier da outra função. Isto acontece pois, devido ás propriedades de simetria da função Hermitiana, a fase de sua transformada de Fourier é zero ou π . Assim, para o problema (P3) a deconvolução é única quando a função não Hermitiana satisfaz a condição de unicidade.

Como o problema de recuperação de fase é um caso particular da deconvolução de funções, onde $g(x, y) = f(x, y) \star f^*(-x, -y) = \mathcal{F}(F^*(u, v)F(u, v))$ com F sendo a transformada de Fourier de f e \star e \star denotando, respectivamente, a convolução e a conjugação complexa, este é mais um caso onde estão definidas as condições de unicidade de solução para o problema de deconvolução.

Barakat e Newsam [3] mostraram que o problema da fase (P1 e P2) é localmente mal condicionado no que se refere a condição (iii).

Ambos os problemas, da recuperação da fase ou da amplitude, tem aplicações muito importantes. A seguir descreveremos algumas delas.

1.2 Aplicações

1.2.1 Astrofísica, Interferometria Especular

Um dos problemas reais mais importantes que são modelados pelo problema matemático da recuperação da fase, e que tem sido nossa principal fonte de referência por se tratar de reconstrução de imagens, é a chamada interferometria especular (tradução livre de 'speckle interferometry', sendo que 'speckle' poderia ser traduzido como 'salpicado', já que se trata de um tipo particular de distorção) (IE). Desde a época de Newton, até a década dos 70, a atmosfera terrestre tem sido reconhecida como uma limitação insuperável para a resolução de telescópios astronômicos baseados na terra. Este inconveniente foi superado com a introdução por Labeyrie [21] em 1970 da técnica chamada de IE estelar.

Para cada imagem instantânea gravada, a equação que se aplica é

$$i(x,y) = o(x,y) \star p(x,y), \tag{1.6}$$

onde i(x, y) representa a intensidade instantânea da imagem, o(x, y) a intensidade do objeto que queremos reconstruir, e p(x, y) a função de espalhamento pontual (PSF) da atmosfera/telescópio. O problema central consiste então em conseguir uma boa aproximação da função de PSF; como isto não era possível antes da introdução da IE, era necessário resolver um problema de deconvolução cega completamente indeterminado. A IE é uma técnica que consegue estimar, no lugar de p(x, y), uma média da função de transferência T(u, v), que é a TF de p(x, y). Obtém-se uma equação da forma

$$\langle |I(u,v)|^2 \rangle = |O(u,v)|^2 \langle |T(u,v)|^2 \rangle, \tag{1.7}$$

onde $I \in O$ denotam as TF de $i \in o$ respectivamente $e \langle , \rangle$ as médias sobre um conjunto de imagens simultâneas. Da equação anterior se deduz que somente é possível obter com precisão adequada a amplitude da TF do objeto e assim ocorre o problema da recuperação da fase (P2). Uma descrição com mais detalhes das características da IE podem ser encontradas em [4] e em [5] (Capítulo 4). O problema (P4) aparece também em

astrofísica quando é necessário recuperar um sinal (ou imagem) que passou por alguma forma de distorção, isto é, o sinal é a convolução de uma função de distorção com o sinal que deve ser recuperado e, por exemplo, a fase da transformada de Fourier de tal função é nula (isso ocorre em casos de longa exposição a turbulência atmosférica). Então a fase do sinal não é distorcida e portanto a partir da convolução obtém-se a fase da transformada de Fourier do sinal a ser recuperado. É preciso então recuperar a amplitude do sinal que é o problema (P4).

1.2.2 Cristalografia

O problema (P2) ocorre em cristalografia onde o raio X é utilizado para determinar a estrutura das moléculas de um cristal, isto é, determinar a posição dos átomos na molécula ([31], [44]). Para determinar a fase de uma onda é preciso ser capaz de monitorar a onda completamente e precisamente. Como raios X tem uma frequência muito alta ($\approx 10^{18} HZ$) isso implica que um instrumento capaz de medir a fase de um raio X deve ser capaz de detectar mudanças num campo elétrico ocorrendo em tempos da ordem de 10^{-18} segundos. Tal instrumento não existe, logo somente a amplitude da onda difratada pode ser medida. Mesmo no caso da amplitude, como uma molécula é extremamente pequena é necessário utilizar uma amostra do cristal contendo muitas moléculas identicamente orientadas e igualmente espaçadas, para obter amplitudes difratadas possíveis de serem medidas.

O campo de espalhamento dos raios X é aproximadamente a TF da densidade eletrônica f(x) no cristal. Portanto em condições normais $f(x) \ge 0$. Como comentado anteriormente o problema da fase aparece porque somente a intensidade dos raios difratados (proporcional a $|F(u)|^2$) pode ser medida.

Um cristal pode ser descrito, em termos de uma célula elementar que é repetida por múltiplas translações ao longo de três vetores elementares no espaço tridimensional. Se cada célula é referenciada como um ponto, o conjunto infinito dos pontos assim formados é chamado de rede. A densidade eletrônica no cristal pode ser escrita como

$$f(x) = e(x) \star l(x) = e(x) \star \sum_{n = -\infty}^{+\infty} \delta(x - z_n)$$
(1.8)

onde \star denota convolução, l(x) a rede, os z_n são os pontos da rede e e(x) denota a densidade eletrônica numa única célula elementar (e zero em outro lugar). Existem vários tipos de redes, mas, para facilitar a exposição consideramos a rede primitiva ortorrômbica em que os vetores elementares são mutuamente ortogonais e existe uma molécula por célula elementar. A expressão anterior para a densidade eletrônica pode ser escrita da seguinte forma

$$f(x_1, x_2, x_3) = e(x_1, x_2, x_3) \star \sum_{m, n, p=0}^{+\infty} \delta(x_1 - ma) \delta(x_2 - nb) \delta(x_3 - pc), \quad (1.9)$$

onde $x = (x_1, x_2, x_3)$ a,b e c são os comprimentos dos vetores elementares e δ denota a função delta ou distribuição. A célula elementar pode ter simetrias internas chamadas simetrias de grupo espacial (space group symmetries), mas para evitar detalhes técnicos não importantes aqui consideramos uma célula elementar sem simetrias. Quando, por exemplo, o conjunto de átomos moléculas ou conteúdo de uma célula elementar é centrosimétrico e a origem se encontra sobre o centro de simetria a parte imaginária de F(u) é nula e então o problema da fase se reduz a uma decisão de sinal isto é F(u) = |F(u)| ou F(u) = -|F(u)|. Se a célula elementar contém N átomos então a densidade eletrônica na célula elementar é dada por

$$e(x_1, x_2, x_3) = \sum_{n=1}^{N} e_n(x_1 - z_1, x_2 - z_2, x_3 - z_3), \qquad (1.10)$$

onde $e_n(x_1, x_2, x_3)$ é a densidade eletrônica no n-ésimo átomo e (z_1, z_2, z_3) a posição do nésimo átomo. Como a densidade eletrônica está concentrada em torno do núcleo atômico, e(x) salta em posições atômicas (propriedade conhecida como atômicidade) logo determinada a função densidade eletrônica e(x) é possível inferir as posições dos átomos. O tipo de átomo pode ser determinado a partir da amplitude de e(x), que é proporcional ao número atômico.

Em cristalografia a função a ser reconstruída é e(x) e não f(x) que é de suporte infinito. Como os raios X são difratados por todo cristal, a amplitude medida é $|F(u)| = |F(u_1, u_2, u_3)|$, onde pelas equações (1.1) e (1.7)

$$F(u_1, u_2, u_3) = E(u_1, u_2, u_3). \sum_{h,k,l=-\infty}^{+\infty} \delta(u_1 - h/a) \delta(u_2 - k/b) \delta(u_3 - l/c)$$
(1.11)

e E(u) é igual á TF de e(x), o padrão de difração será obtido de uma única célula elementar. De (1.9) e (1.11) tem-se que as medidas F(u) são portanto a TF da densidade eletrônica de uma única célula amostrada em pontos (h/a, k/b, l/c). Em cristalografia o espaço de Fourier é chamado espaço recíproco e os pontos (h/a, k/b, l/c) formam a rede recíproca. É convencional denotar as amostras de $F(u_1, u_2, u_3)$, chamadas de fatores de estrutura, por F_{hkl} , isto é,

$$F_{hkl} = F(h/a, k/b, l/c) = E(h/a, k/b, l/c) = E(u)L(u)$$
(1.12)

onde L(u) denota a rede recíproca.

É claro que

$$F_{hkl} = \int_0^c \int_0^b \int_0^a f(x_1, x_2, x_3) \exp[i2\pi (hx_1/a + kx_2/b + lx_3/c] dx_1 dx_2 dx_3$$
(1.13)

ou

$$F_v = \int_V f(x) \exp(i2\pi v) dx \tag{1.14}$$

onde v = (h/a, k/b, l/c) e V é o volume da célula elementar.

É importante observar na equação (1.11) que a intensidade devida a célula elementar $(|E(u)|^2)$, que é uma função contínua, é submetida a uma amostragem (rede recíproca). Em outras palavras, a repetição do modelo atômico da célula elementar tem por conseqüência uma perda de informação no espaço recíproco já que então só é possível observar o quadrado do módulo da transformada de Fourier em pontos discretos da rede e não em todo o espaço recíproco . Nos cristais reais (não ideais) encontra-se intensidade apreciável

fora desses pontos (picos de Bragg), devido ao fato da rede não ser perfeita, o cristal ser infinito, os átomos estarem vibrando por efeito térmico, existirem átomos fora do lugar correto ou substituidos por átomos diferentes ou por outros defeitos.

Se a fase de F_{hkl} também pudesse ser medida não haveria problema pois as amostras são separadas pelo intervalo de Nyquist e são portanto suficientes para reconstruir continuamente $E(u_1, u_2, u_3)$ pelo teorema da amostragem. Como o suporte da autocorrelação $(\mathcal{F}^{-1}(|F(u)|^2))$ em qualquer direção é duas vezes o suporte de e(x) na mesma direção o intervalo de Nyquist para $|E(u)|^2$ é portanto metade que para E(u) e assim $|E(u)|^2$ contínua não pode ser reconstruída com os dados amostrados. Esta particularidade no problema de recuperação de fase em cristalografia foi observada pela primeira vez por Sayre [47]. Segue disso que a autocorrelação de e(x) não pode ser construida utilizando somente $|F_{hkl}|$. O que pode ser obtida é a autocorrelação de f(x) que é chamada de função de Patterson de e(x),

$$P(x) = (1/V) \sum_{v} |F_{v}|^{2} \exp(i2\pi v . x).$$
(1.15)

Pelas equações (1.14) e (1.15),

$$P(x_1, x_2, x_3) = \sum_{-\infty}^{+\infty} A(x_1 - ha, x_2 - kb, x_3 - lc)$$
(1.16)

onde A(x) é a autocorrelação de e(x). Como A(x) tem duas vezes a extensão de e(x) a autocorrelação em (1.16), se sobrepõe. Tal dificuldade é interpretada como um "aliasing problem" no espaço da autocorrelação. Como e(x) é atomística, assim é P(x), e picos em P(x) ocorrem em posições que correspondem a vetores interatômicos na estrutura. Em princípio, P(x) pode ser usada para se obter e(x), mas em alguns casos isso é praticamente impossível.

No caso de cristalografia macromolecular, que determina a estrutura de moléculas com mais de 500 átomos, os métodos diretos não são efetivos pelas seguintes razões [31]: - o grande número de átomos resulta numa grande variedade de fases prováveis. -Os dados de difração são medidos usualmente somente para resolução moderada de modo que a propriedade de atomicidade em que os métodos diretos são baseados não é refletida nos dados.

Todos os procedimentos utilizados em cristalografia macromolecular utilizam alguma informação auxiliar que sempre aproxima a fase a ser calculada. Esta informação é obtida com uma variedade de técnicas experimentais ou por usar uma estrutura similar conhecida. A estimativa obtida juntando as intensidades medidas e a aproximação da fase é utilizada como estimativa inicial para a aplicação dos algoritmos iterativos propostos por Fienup que descreveremos posteriormente. Assim tais algoritmos funcionam como um processo de refinar a aproximação da densidade eletrônica obtida.

1.3 Panorama desta Tese

Após a apresentação dos problemas neste capítulo introdutório, no próximo capítulo apresentaremos uma descrição dos métodos existentes para resolvê-los. Nos concentramos no problema (P2), o mais difícil pela não convexidade e também o mais importante nas aplicações, mas algumas metodologias são comuns para os outros problemas relacionados. Também nossa referência principal é o caso bidimensional, a recuperação de imagens, e as informações adicionais são essencialmente não negatividade e aquela que provém do suporte da autocorrelação. Descreveremos os métodos para recuperar tipos especiais de objetos, e depois os métodos não iterativos para objetos gerais. Na terceira Seção apresentamos uma descrição detalhada dos métodos iterativos, cujo estudo é o principal objetivo desta dissertação. Damos uma breve explicação do problema da estagnação e das técnicas, previamente existentes, propostas para superá-lo.

Os Capítulos 3, 4, 5, 6 e 7 contém os resultados originais desta tese, produto da nossa pesquisa sobre os métodos iterativos.

O Capítulo 3 descreve alguns novos métodos alternativos para resolver (P2) e (P3), que em alguns casos tem produzido melhores resultados que os métodos anteriores.

As questões relacionadas com a convergência e a estagnação dos métodos para resolver

(P2) nos levaram à necessidade de conseguir uma caracterização da geometria dos conjuntos das possíveis soluções envolvidas, no caso discreto. Essa caracterização é apresentada no Capítulo 4.

Claramente o primeiro passo para garantir o bom comportamento de um algoritmo é provar que a seqüência definida por ele é limitada. Isto é trivial no caso de alguns algoritmos como o de Gerschberg-Saxton [20] e os algoritmos de Fienup conhecidos por "Error-Reduction" e "Output-Output" [15], mas não é para os outros casos. No Capítulo 5 completamos as demonstrações faltantes de limitação para uma variante do algoritmo híbrido (que é reconhecido universalmente como o mais efetivo) e para aquele conhecido como "Input-Output".

No Capítulo 6 apresentamos uma nova metodologia para superar o problema da estagnação, baseada na caracterização do Capítulo 4. O ponto de partida para esta nova metodologia é completamente novo e possibilita que esta seja independente do tipo de estagnação e das características da solução.

No Capítulo 7 apresentamos parte da experimentação numérica realizada e finalmente no último Capítulo apresentamos algumas conclusões junto com as atuais e futuras direções de pesquisa.

Capítulo 2

MÉTODOS

Descrevemos neste Capítulo os principais métodos existentes para resolver o problema (P2), começando por aqueles que são adequados para alguns tipos especiais de objetos, seguindo por aqueles para objetos mais gerais e completando com mais detalhes, na terceira Seção, com os métodos iterativos que são o principal objeto da nossa pesquisa. Descricões mais detalhadas podem ser encontradas em [5] e [12].

2.1 Métodos para Tipos Especiais de Objetos

2.1.1 Holografia

Quando o objeto inclue uma função delta como componente, suficientemente separada do resto do objeto, então o objeto pode ser encontrado como um termo na autocorrelação. Embora a holografia tenha sido originalmente criada para aplicações empregando luz coerente o mesmo princípio é aplicado para luz incoerente se são dadas as amplitudes da transformada de Fourier do objeto [12].

2.1.2 Difference Fourier Syntheses

Alguns métodos utilizados em cristalografia podem ser usados para aplicações em astronomia, por exemplo, o que segue.

Suponha que o objeto consiste numa coleção de pontos como estrelas e se tem uma estimativa do objeto que contém algumas mas não todas as estrelas. Cria-se uma nova estimativa:

- (1) Calculando a TF da estimativa do objeto;
- (2) Trocando a amplitude da TF pela verdadeira (dados);
- (3) Calculando a TF inversa;

A estimativa resultante dos passos acima conterá as estrelas que faltam na estimativa inicial, mas com metade de seus valores verdadeiros. Assim por tomar a diferença entre a nova estimativa e a inicial pode-se identificar a localização e o brilho relativo das estrelas que faltam e utilizar a informação para uma nova estimativa do objeto. Este algoritmo deve ser aplicado iterativamente porque são produzidos termos estranhos junto com as estrelas faltantes desejadas. Este método tem bons resultados para objetos pontuais mas não para objetos extendidos [12].

2.1.3 Produto de Autocorrelações

Se o objeto consiste de uma coleção de estrelas, então sua autocorrelação $r_f(x, y)$ também consiste de um número de termos como pontos. Primeiro computa-se o produto $r_f(x - x_0, y - y_0)r_f(x, y)$, onde $r_f(x_0, y_0)$ é não nula. Em seguida computa-se o produto triplo $r_f(x - x_1, y - y_1)r_f(x - x_0, y - y_0)r_f(x, y)$, onde $r_f(x_1 - x_0, y_1 - y_0)r_f(x_1, y_1)$ é não nula. Se o objeto satisfaz algumas condições de não redundância, então o suporte do produto triplo é igual ao suporte do objeto (ou do objeto rodado por 180 graus). Mais ainda, os valores do objeto podem ser reconstruídos por algumas equações simples envolvendo tomar a raíz cúbica de pontos no produto triplo [18]. As condições de não redundância que o objeto deve satisfazer são:

(i) os vetores de separação entre três pares de seis pontos distintos não devem somar zero.

(ii) os vetores de separação entre dois pares diferentes de pontos não devem ser iguais;

(ii) o vetor de separação entre um par de pontos não deve ser igual ao dobro do vetor de separação entre outro par de pontos.

Em geral uma coleção de estrelas distribuída aleatóriamente satisfaz as condições de não redundância, a menos que o número de estrelas na imagem seja muito grande.

2.2 Métodos para Objetos Gerais

Conhecer a amplitude da transformada de Fourier de uma função f é equivalente a conhecer a autocorrelação de f, pois $\mathcal{F}^{-1}(|F(u,v)|^2) = \mathcal{F}^{-1}(F(u,v).\overline{F(u,v)}) = f(x,y) \star \overline{f(-x,-y)} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi_1,\xi_2) \cdot f(x+\xi_1,y+\xi_2) d\xi.$

Se, por exemplo, f é uma função real de duas variáveis e forem consideradas N^2 amostras, o sistema gerado pela autocorrelação discreta é sobredeterminado com N^2 variáveis e M^2 equações (uma para cada valor de |F(u, v)|). Considerando somente as equações não redundantes serão $N^2 + (N - 1)^2$ equações (ver [37]).

Uma forma clássica de resolver tal sistema é definir um erro métrico e utilizar o método de Newton-Raphson ou um método que utilize o gradiente para procurar uma solução, que é uma estimativa para a qual o erro métrico é zero. O erro métrico poderia ser, por exemplo,

$$E_r = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} [r_g(x,y) - r_f(x,y)]^2$$
(2.1)

onde $r_g(x, y)$ é a autocorrelação de g(x, y), uma estimativa de f(x, y). Se $E_r = 0$ para uma g(x, y) que satisfaz as restrições do domínio do objeto então g(x, y) é uma solução. Tal solução não será necessariamente igual a f(x, y). O erro métrico E_r pode ser visto como uma função num espaço N^2 -dimensional onde os valores de g(x, y) são as coordenadas

do espaço. A coleção dos N^2 valores de g(x, y) pode ser considerada como um vetor em tal espaço. O problema é encontrar, dentro de um subconjunto do espaço onde g(x, y)satisfaz as restrições de não negatividade e de suporte, o ponto em que E_r é um mínimo global ($E_r = 0$ se não houver ruído). Pode-se procurar o mínimo global computando o gradiente de E_r numa estimativa $g_k(x, y)$ e usar esta informação para encontrar uma nova estimativa $g_{k+1}(x, y)$ onde E_r é menor. Continua-se iterativamente até a solução global ser encontrada ou até ocorrer estagnação em algum mínimo local.

2.2.1 Newton-Raphson

O método de Newton-Raphson envolve a linearização local do problema por expressar $r_g(x, y)$ na equação (2.1) pela expansão em série de Taylor em torno da estimativa $g_k(x, y)$:

$$g(x,y) = g_k(x,y) + \Delta g(x,y), \qquad (2.2)$$

e tomando somente o termo de primeira ordem:

$$r_g(x,y) \simeq r_{g_k}(x,y) + \sum_{x'} \sum_{y'} \frac{\partial r_g(x,y)}{\partial g(x',y')} \Delta g(x',y').$$

$$(2.3)$$

Substituindo $r_g(x, y)$ na equação de E_r , E_r é minimizado por considerar iguais a zero as derivadas parciais de E_r com respeito a $\Delta g(x_0, y_0)$ para cada valor de (x_0, y_0) . Isto dá um sistema de N^2 equações em N^2 variáveis $(\Delta g's)$, que pode ser resolvida pela inversão de uma N^2 por N^2 matriz. Uma nova estimativa é dada pela equação (2.2) e o procedimento é repetido. A inversão da matriz é feita com aproximadamente N^6 operações, fazendo com que este método seja muito caro computacionalmente mesmo que poucas iterações sejam necessarias [12]. Também a probabilidade de convergência a um ponto que não é solução é muito alta.

2.2.2 Gradientes

Embora requerendo muito mais iterações que o método de Newton-Raphson os algoritmos que envolvem gradientes são atrativos pois utilizam um número bem menor de operações. Se o cálculo do gradiente é feito utilizando diferenças finitas o tempo envolvido é muito grande (são necessarias uma ou duas vezes N^2 transformadas rápidas de Fourier), mas se o erro métrico é definido no domínio de Fourier por

$$E_F^2 = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} [|G(u,v)| - |F(u,v)|]^2$$
(2.4)

onde G(u, v) é a tranformada de Fourier de g(x, y), é possível mostrar que o gradiente pode ser encontrado a partir dos três primeiros passos dos algoritmos iterativos que descreveremos na próxima seção.

Em geral a k-ésima iteração de um método de busca pelo gradiente é dada por

$$g_k''(x,y) = g_k(x,y) + h_k d_k(x,y)$$
(2.5)

onde $d_k(x, y)$ é a direção que se toma para encontrar a nova estimativa a partir de $g_k(x, y)$ e h_k é o comprimento do passo nesta direção. Para o método de máxima descida as coordenadas de $d_k(x, y)$ são o negativo das derivadas parciais de E_F^2 em relação a g(x, y). Para o método dos gradientes conjugados, $d_k(x, y)$ é uma combinação linear do negativo do gradiente e de $d_{k-1}(x, y)$. Após este passo considera-se a estimativa mais próxima de $g''_k(x, y)$ que satisfaz as restrições do domínio do objeto para ser $g_{k+1}(x, y)$ [12].

2.2.3 Método do teorema da amostragem

A utilização do teorema da amostragem [7] está vinculada à condição do objeto ser de dimensão finita. Se o objeto é zero fora de um quadrado de comprimento L centrado na origem, então o teorema da amostragem leva a

$$F(u,v) = \sum_{p} \sum_{q} F(p\Delta u, q\Delta v) \operatorname{sinc}\left(\frac{u - p\Delta u}{\Delta u}\right) \operatorname{sinc}\left(\frac{v - q\Delta v}{\Delta v}\right), \quad (2.6)$$

onde $\Delta u = \Delta v = L^{-1}$. No método do teorema da amostragem, as amostras $F(p\Delta u, q\Delta v)$ são estimadas por $|F(p\Delta u, q\Delta v)| exp[i\phi(p\Delta u, q\Delta v)]$, onde ϕ é uma estimativa da fase desconhecida. Isto é incluido no lado direito da equação, que é calculada na metade de suas coordenadas inteiras, e o quadrado da amplitude é também calculado. Como uma amostragem adequada de $|F(u,v)|^2$ contém duas vezes o número de amostras requeridas para F(u, v) e, é assumido que $|F(u, v)|^2$ foi adequadamente amostrada, os quadrados das amplitudes calculados podem ser comparados com os valores de $|F(u,v)|^2$. Isto leva a um sistema de N^2 equações não lineares a N^2 variáveis, que pode ser resolvido, por exemplo, pelo método de Newton-Raphson. A diferença deste método para o de Newton-Raphson descrito anteriormente está somente nas equações, assim a complexidade computacional é a mesma. Algumas versões um pouco mais eficientes do método do teorema da amostragem estão descritas em [12].

2.2.4 Simulated Annealing

Lawton, Morrison e outros tem tentado "simulated annealing" para resolver o problema de recuperação de fase. Para uma dada função objetivo H [26] o algoritmo de Lawton e Morrison é basicamente o seguinte:

1) Escolher uma estratégia para selecionar candidatos θ'_{k+1} para θ_{k+1} , dado o corrente valor de θ_k ;

- 2) Selectionar $\beta > 0$ e escolher θ_1 ;
- 3.1) Usar a estratégia para calcular θ'_{k+1} ;
- 3.2) Calcular $\Delta H = H(\theta'_{k+1}) H(\theta_k)$
- 3.3) Se $\Delta H \leq 0$ fazer $\theta_{k+1} = \theta'_{k+1}$
- 3.4) Se $\Delta H > 0$ fazer

$$\theta_{k+1} = \begin{cases} \theta'_{k+1} & \text{com probabilidade } e^{-\beta \Delta H} \\ \theta_k & \text{com probabilidade } 1 - e^{-\beta \Delta H} \end{cases}$$

4) Incrementar β e aplicar o passo 3 até obter um resultado satisfatório.

2.2.5 Levenberg-Marquardt

Nieto-Vesperinas ([37]) resolveu o sistema sobredeterminado gerado pela autocorrelação aplicando a modificação (feita por Levenberg-Marquardt) do método de Gauss-Newton ao problema de quadrados mínimos associado. O principal problema deste método é sua convergência para mínimos locais não globais (que são as soluções do problema). Tentando evitar este problema, durante nosso trabalho de pesquisa, utilizamos a extensão não linear do método de Kaczmarz (Newton aplicado linha a linha), com uma ordem aleatória das linhas, para resolver o problema. Nosso objetivo não foi alcançado, isto é, ainda ocorreram pontos fixos diferentes da solução.

2.3 Métodos Iterativos

2.3.1 Métodos iterativos para recuperar a fase da transformada de Fourier

Os algoritmos iterativos tem sido os métodos com maior sucesso para a reconstrução de objetos não negativos a partir da amplitude de sua TF. O motivo disto é que eles tratam com tipos de objetos mais gerais, usam todos os dados e restrições avaliáveis, não são altamente sensíveis a ruidos e não são computacionalmente caros como a maior parte dos outros métodos. Em geral, quando se quer recuperar imagens, os métodos iterativos envolvem de alguma forma projeções sobre conjuntos convexos ou não convexos. O objetivo fundamental dos algoritmos que utilizam projeções é o de forçar o sinal ou imagem a pertencer a m conjuntos $C_1, C_2, ..., C_m$ definidos, respectivamente, pelas m propriedades do sinal ou imagem. Por exemplo, se a restrição é $f \geq 0$ então f terá que pertencer a C_p

(conjunto de todas as funções reais não negativas). Associado a cada C_i está o operador projeção P_i (i = 1, ...m). Para todo conjunto fechado (não necessariamente convexo) dizse que $g \approx P_i h$ é a projeção de h sobre C_i se $||g - h|| = min_{y \in C_i}||y - h||$ (i = 1, 2, ..., m). Esta projeção existe e é única se C_i é fechado e convexo ([52]). Quando C_i não é convexo não está garantida a existência de uma projeção, porém é difícil conjecturar uma situação realista onde tal projeção não exista. Pode existir um conjunto de pontos que satisfazem a definição de projeção, porém, no nosso caso de interesse, é possível encontrar um método de escolher um desses pontos unicamente, usualmente por satisfazer outra condição.

Por comodidade descreveremos alguns métodos iterativos considerando recuperação de funções de uma variável. Os outros casos são análogos.

O método mais conhecido para resolver o problema (P1), descrito na introdução, é o algoritmo proposto por **Gerchberg e Saxton** em 1972 [20] que partindo de uma estimativa inicial do sinal obtém, através de projeções e de forma iterativa, novas estimativas utilizando os seguintes sub-passos:

(1) Cálculo da transformada de Fourier da estimativa;

(2) Projeção no domínio de Fourier do ponto obtido em (1) sobre o conjunto (não convexo)
 das funções cujas amplitudes coincidem com a amplitude da transformada de Fourier da
 função a ser recuperada (dados);

(3) Cálculo da transformada de Fourier inversa da projeção obtida em (2);

(4) Projeção do ponto obtido em (3) sobre o conjunto (também não convexo) das funções com amplitudes iguais às da função a ser recuperada (dados);

Observamos que apesar da não garantia da existência ou unicidade da projeção sobre um conjunto não convexo neste problema é possível utilizá-la pois a fase da transformada de Fourier da k-ésima estimativa define unicamente a projeção. Ela sempre está definida e, excluindo os casos onde há amplitudes nulas, é única. Mesmo no caso em que ocorrem amplitudes nulas é possível definir a projeção de maneira única. Assim, se a função a ser recuperada é $f(x) = |f(x)|\exp[i\eta(x)]$ com transformada de Fourier $F(u) = |F(u)|\exp[i\psi(u)]$, defini-se a k-ésima iteração do algoritmo de Gerchberg-Saxton da seguinte forma:

$$\begin{split} G_k(u) &= |G_k(u)| \exp[i\phi_k(u)] = \mathcal{F}[g_k(x)] \\ G'_k(u) &= |F(u)| \exp[i\phi_k(u)] \\ g'_k(x) &= |g'_k(x)| \exp[i\theta'_k(x)] = \mathcal{F}^{-1}[G'_k(u)] \\ g_{k+1}(x) &= |f(x)| \exp[i\theta_{k+1}(x)] = |f(x)| \exp[i\theta'_k(x)] \\ \text{onde } g_k, \ \theta_k, \ G'_k \ e \ \phi_k \ \text{são, respectivamente, estimativas de } f, \ \eta \ , \ \mathbf{F} \ e \ \psi. \end{split}$$

Em geral, o algoritmo é iniciado com a utilização de um vetor $g_0(x)$ aleatório ou de uma fase $\phi_0(u)$ aleatória. Como G'_k é a projeção de G_k sobre o conjunto das funções cujas amplitudes coincidem com as amplitudes de F (que denotaremos por C_2), o quadrado da distância de G_k a C_2 é (no caso finito e pensando as funções como vetores N-dimensionais)

$$d^{2}(G_{k}, C_{2}) = N^{-1} \sum_{u=1}^{N} |G_{k}(u) - G_{k}'(u)|^{2}$$
(2.7)

que pode ser escrito como

$$d^{2}(G_{k}, C_{2}) = N^{-1} \sum_{u=1}^{N} [|G_{k}(u)| - |F(u)|]^{2}.$$
(2.8)

Analogamente, no domínio do tempo, tem-se que o quadrado da distância de g'_k ao conjunto das funções com amplitudes coincidindo com as da função a ser recuperada (que chamaremos de C_1) é

$$d^{2}(g'_{k}, C_{1}) = \sum_{x=1}^{N} [|f(x)| - |g'_{k}|]^{2}.$$
(2.9)

Fienup [17] demonstrou que $d^2(G_k, C_2)$ é não crescente com as iterações do algoritmo. Além disso adaptou o algoritmo de Gerchberg-Saxton para o problema P2 e provou que neste caso $d^2(G_k, C_2)$ também é não crescente. Além dessa propriedade não há nada provado sobre a convergência desses algoritmos.

A adaptação do algoritmo de Gerchberg e Saxton para o problema P2, que foi chamada por Fienup de algoritmo "Error-Reduction" (ER), difere do original somente no quarto sub-passo que é substituído pela projeção de g'_k sobre o conjunto definido pela condição convexa sobre f, isto é, tal sub-passo é dado por

$$g_{k+1}(x) = \begin{cases} g'_k(x) & \text{se } x \in \gamma \\ 0 & \text{se } x \notin \gamma \end{cases}$$
(2.10)

onde γ é o conjunto dos pontos em que $g_k'(x)$ satisfaz as restrições do objeto. Aqui $C_1=\gamma$ e

$$d^{2}(g'_{k}, C_{1}) = \sum_{x \notin \gamma} [g'_{k}(x)]^{2}$$
(2.11)

Em Astronomia, por exemplo, as restrições no domínio do objeto são a não negatividade do objeto e uma restrição de suporte (usualmente perdida). O diâmetro do objeto pode ser encontrado pois é exatamente metade do diâmetro da autocorrelação, mas o suporte exato do objeto não pode, em geral, ser unicamente determinado a partir do Observamos que tanto para o algoritmo Gerchberg-Saxton suporte da autocorrelação. quanto para o ER nenhuma das distâncias definidas acima é a distância da estimativa obtida pelo algoritmo à solução do problema. A solução pertence a $C_1 \cap \mathcal{F}^{-1}[C_2]$. Assim, embora $g_k \in C_1$ e $g'_k \in C_2$, o não crescimento de $d^2(G_k, C_2)$ ou de $d^2((g'_k, C_1)$ não significa aproximação da solução. A única garantia que se tem é que quando uma dessas distâncias é nula obteve-se uma solução (g_k ou g'_k). Embora $d^2(G_k, C_2)$ seja não crescente com as iterações deste algoritmo esta pode decrescer rapidamente nas primeiras iterações e depois muito lentamente ou até estagnar sem que o algoritmo tenha encontrado a solução. No final deste capítulo, descreveremos algumas formas de estagnação que foram identificadas por Fienup e alguns métodos para sair destes pontos foram propostos em [19]. Não há uma caracterização matemática destas estagnações.

Tentando evitar estagnações, Fienup propôs uma classe de outros algoritmos. O primeiro, denominado Input-Output (IO) difere dos algoritmos anteriores no quarto subpasso que é definido por

$$g_{k+1}(x) = \begin{cases} g_k(x) & \text{se } x \in \gamma \\ g_k(x) - \beta g'_k(x) & \text{se } x \notin \gamma \end{cases}$$
(2.12)

onde β é um parâmetro escolhido adequadamente e γ é o conjunto dos pontos em que $g'_k(x)$ satisfaz as restrições do objeto. Na prática o algoritmo IO tem convergido mais

rapidamente para ambos os problemas (P1) e (P2).

Outro algoritmo proposto é o denominado "Output-Output" (OO) derivado do ER pela relaxação da projeção no quarto sub-passo, isto é,

$$g_{k+1}(x) = \begin{cases} g'_k(x) & \text{se } x \in \gamma \\ g'_k(x) - \beta g'_k(x) & \text{se } x \notin \gamma \end{cases}$$
(2.13)

onde γ é o conjunto dos pontos em que $g'_k(x)$ satisfaz as restrições do objeto e β é o parâmetro de relaxação. Este algoritmo costuma se tornar bastante lento mesmo quando ainda está distante de uma solução. Na tentativa de evitar as estagnações do OO, Fienup combinou a primeira linha da equação (2.13) com a segunda linha da equação (2.12) para gerar o algoritmo conhecido como Input-Output Híbrido (IOH) onde

$$g_{k+1}(x) = \begin{cases} g'_k(x) & \text{se } x \in \gamma \\ g_k(x) - \beta g'_k(x) & \text{se } x \notin \gamma \end{cases}$$
(2.14)

onde β é um parâmetro escolhido adequadamente e γ é o conjunto dos pontos em que $g'_k(x)$ satisfaz as restrições do objeto. Nestes três últimos algoritmos as estimativas g_k não satisfazem as restrições do domínio do objeto. Assim medir $d^2(G_k, C_2)$ não faz sentido. Quando $d^2(g'_k, C_1)$ é zero então g'_k é uma solução.

Na prática o algoritmo IOH é o melhor da classe dos algoritmos propostos por Fienup embora **não exista nenhuma prova da convergência do mesmo.** Mesmo tendo sentido medir $d^2(g'_k, C_1)$ esta nem sempre concorda com a qualidade da imagem. Por esta razão, costuma-se fazer uma combinação de iterações do híbrido seguida de iterações do ER. O algoritmo ER, aparentemente faz o papel de refinar uma aproximação da solução obtida pelo híbrido.

O algoritmo das projeções generalizadas, que descreveremos a seguir, pode ser utilizado tanto para resolver os problemas de recuperação de fase como para resolver o problema de recuperação de amplitude P3.

Suponhamos que existam m condições sobre a imagem a ser reconstruída dando origem a m conjuntos (convexos ou não). A iteração do algoritmo das projeções generalizadas é dada por

$$g_{k+1} = T_1 T_2 \dots T_m g_k \tag{2.15}$$

onde $T_i = I + \lambda_i (P_i - I)$ com I sendo o operador identidade e λ_i uma constante chamada parâmetro de relaxação.

A demonstração de convergência da seqüência(2.15) é bastante conhecida (as referências correspondentes podem ser encontradas em [52]) tendo como hipóteses que:

(i) C_i (i=1,...m) sejam fechados com intersecção não vazia , isto é, $\cap_{i=1}^m C_i \neq \phi$

(ii) $0 < \lambda < 2$ para i = 1, ...m.

Como já observamos, na recuperação da fase existem conjuntos que não são convexos e neste caso não há uma prova da convergência do algoritmo das projeções generalizadas. Porém, Levi e Stark [29] obtiveram algumas conclusões interessantes para o caso em que m = 2 (que não é uma restrição muito séria pois sempre é possível combinar duas ou mais restrições para formar um único conjunto).

Consideremos a função

$$J(g_k) = ||P_1g_k - g_k|| + ||P_2g_k - g_k||$$
(2.16)

onde P_1 e P_2 são, respectivamente, as projeções sobre C_1 e C_2 . Levi e Stark mostraram que, com determinadas condições sobre os parâmetros de relaxação J é não crescente com as iterações do algoritmo dado pela equação (2.15). Mais ainda, se $J(g_k) = 0$ então g_k é uma solução. As condições citadas se encontram no Capítulo 3 desta tese.

Os algoritmos ER e OO são casos particulares do algoritmo das projeções generalizadas, bastando considerar $\lambda_1 = \lambda_2 = 1$ para obter o primeiro e $\lambda_2 = 1$ para o segundo. No caso do algoritmo ER a função J coincide com a distância $d^2(g_k, C_2)$ que Fienup mostrou ser não crescente com as iterações do algoritmo, pois $||P_1g_k - g_k|| = 0$.

2.3.2 Métodos Iterativos para Recuperar a Amplitudede da Transformada de Fourier

Hayes [22] propôs um método para recuperar uma seqüênciax(n) *m*-dimensional, com suporte finito e com Z-transformada sem fatores simétricos, dada a fase de sua transformada de Fourier (neste caso o problema P3 tem solução única). Tal método envolve encontrar a solução de um sistema de equações lineares que provém da definição da fase da transformada de Fourier de x(n). Na prática a aplicação deste método é limitada pelas dificuldades computacionais para encontrar a solução do sistema de equações lineares. Um método iterativo que é utilizado para resolver o problema de recuperação de amplitude é devido a Oppenheim, Hayes e Lim [23]. Este método é similar aos algoritmos de Fienup para recuperação de fase. Mais especificamente, para g_k estimativa do sinal a ser recuperado, o algoritmo efetua os seguintes passos:

(1) Cálculo da transformada de Fourier, G_k , de g_k ;

(2) Troca da fase $\psi_k(u)$ de G_k pela fase $\phi(u)$ dos dados gerando uma estimativa $G'_k(u)$;

(3) Cálculo da transformada de Fourier inversa, g'_k , de $G'_k(u)$;

(4) Formação de uma nova estimativa, g_{k+1} , impondo à g_k as restrições do sinal.

Matematicamente para a k-ésima iteração os quatro passos são:

$$G_{k}(u) = |G_{k}(u)|rmexp[i\psi_{k}(u)] = \mathcal{F}[g_{k}(x)];$$

$$G'_{k}(u) = |G_{k}(u)|rmexp[i\phi(u)];$$

$$g'_{k}(x) = \mathcal{F}^{-1}[G'_{k}(u)];$$

$$g_{k+1}(x) = \begin{cases} g'_{k}(x) & \text{se } x \notin \gamma, x \neq 0 \\ \alpha & \text{se } x = 0 \\ 0 & \text{em outro caso} \end{cases}$$

onde γ é o conjunto dos pontos no interior do suporte da imagem e α é uma constante arbitrária.
2.3.3 Métodos Iterativos para Deconvolução de Funções

O primeiro método iterativo para deconvolução das funções f(x) e h(x) da equação (1.5) que descreveremos aqui é devido a Ayers e Dainty [2]. Este método admite que se tenha alguma informação sobre as funções f(x) e h(x), por exemplo, que são não negativas, e, utilizando o fato de que a TF de (1.5) é dada por

$$G(u) = F(u)H(u), \qquad (2.17)$$

consiste dos seguintes passos:

(1) Cálculo da transformada de Fourier, $F_k^1(u)$, da estimativa $f_k^1(x)$ de f(x);

(2) Inversão de $F_k^1(u)$ para formar um filtro inverso que multiplicado por G(u) gerará uma estimativa $H_k(u)$ para H(u) ($H_k(u) = G(u)/F_k^1(u)$);

(3) Cálculo da transformada de Fourier inversa, $h_k(x)$, de $H_k(u)$;

(4) Projeção de $h_k(x)$ sobre o conjuntos das funções que satisfazem as restrições de h(x), gerando uma nova estimativa $h_k^1(x)$;

(5) Cálculo da transformada de Fourier, $H_k^1(u)$, de $h_k^1(x)$;

(6) Inversão de $H_k^1(u)$ para obter outro filtro inverso que multiplicado por G(u) gerará uma estimativa $F_k(u)$ de F(u);

(7) Cálculo da transformada de Fourier inversa, $f_k(x)$, de $F_k(u)$;

(8) Projeção de $f_k(x)$ sobre o conjunto das funções que satisfazem as restrições sobre f(x), gerando uma nova estimativa $f_{k+1}^1(x)$;

Neste método ocorrem dois problemas:

a) A definição do filtro inverso nas regiões onde $F_k^1(u)$ ou $H_k^1(u)$ tem valores próximos de zero;

b) Zeros em frequências espaciais particulares em uma das funções F(u) ou H(u) resultam em nenhuma informação na convolução;

Para evitar estes problemas, quando as regiões de valores nulos ou baixos na convolução são pequenas e bem definidas, Ayers e Dainty definiram $F_k(u)$ do passo (6) da seguinte forma: Se |G(u)| < nível de ruído então $F_k(u) = F_k^1(u)$ Se $|H_k^1(u)| \ge |G(u)|$ então $F_k(u) = (1 - \beta)F_k^1(u) + \beta \frac{G(u)}{H_k^1(u)};$ Se $|H_k^1(u)| < |G(u)|$ então $\frac{1}{F_k(u)} = \frac{(1 - \beta)}{F_k^1(u)} + \beta \frac{H_k^1(u)}{G(u)};$ onde $0 \le \beta \le 1$.

Em 1992, Lane [28] utilizou o método dos gradientes conjugados para resolver o problema de deconvolução. Minimizou a função $E_c = E_i + E_f$ onde $E_i = \int \int_{\gamma_f} |f(x, y)|^2 dx dy + \int \int_{\gamma_h} |h(x, y)|^2 dx dy \operatorname{com} \gamma_f e \gamma_h$, sendo respectivamente, os conjuntos onde f e h violam as restrições e $E_f = \int \int |G(u, v) - F(u, v)H(u, v)|^2 du dv$ e observou que este método é mais estável que o método de Ayers e Dainty, embora seja computacionalmente mais caro.

Nakajima [34] propôs um método para deconvolver funções no caso particular onde uma das componentes é uma função Hermitiana e a outra é não Hermitiana. Tal método consiste em dois passos. No primeiro passo a fase de Fourier da função não Hermitiana é reconstruída a partir da convolução usando a propriedade de simetria das funções Hermitianas (a transformada de Fourier de uma função Hermitiana é real). No segundo passo a função não Hermitiana é reconstruída a partir da fase obtida no primeiro passo e das restrições de suporte, utilizando o algoritmo para recuperação de amplitude de Hayes, Oppenheim e Lim. Embora o método forneça a solução para uma versão restrita do problema de deconvolução cega, ele pode ser aplicado a algumas situações práticas, como por exemplo, na restauração de imagens borradas. As funções do processo de borrar são Hermitianas e, em geral, as imagens originais são não Hermitianas.

2.3.4 Tipos de Estagnações e Métodos para Sobrepô-las

Como já observamos, em alguns casos os algoritmos iterativos estagnam antes de encontrar a solução. Três dos pontos de estagnação mais comuns, constatados numéricamente, são descritos por Fienup e Wackerman em [19], e caracterizados por: (1) Imagens gêmeas simultâneas; (2) Faixas; e (3) Suporte truncado. Alguns métodos simples de tentar sobrepô-las podem ser: (a) mudar a estimativa inicial; (b) utilizar um suporte diferente por algumas iterações; (c) no caso do algoritmo híbrido, utilizar β maior por algumas iterações [12];

Como ambas f(x, y) e sua imagem gêmea f(-x, -y) tem o mesmo módulo de Fourier, os algoritmos podem reconstruir de forma igual cada uma delas. Quando o suporte é simétrico os algoritmos podem gerar uma imagem parcialmente reconstruída tendo características de ambas. Foi observado numéricamente que em muitos casos os algoritmos estagnam nesta imagem. Um método para sair desta estagnação é utilizar uma redução do suporte por algumas iterações. O suporte temporário deve ser altamente não simétrico e contendo uma ou duas bordas do suporte do objeto mas não a borda oposta. Após umas poucas iterações com o suporte reduzido, a simetria da imagem pode ser suficientemente quebrada e, voltando ao suporte original, mais algumas iterações levarão a f(x, y) ou a sua gêmea.

Os algoritmos também podem estagnar quando g'_k é a imagem procurada mas com um padrão de faixas sobrepostas. Isto pode ser reconhecido pois as faixas se estendem ao exterior do suporte da imagem. O padrão das faixas é aproximadadamente senoidal em uma direção e constante na direção ortogonal. As faixas, em geral, são de baixo contraste e portanto não desagradáveis, mas ocasionalmente são de suficiente contraste para atrapalhar. Com diferentes estimativas iniciais aleatórias, as faixas das imagens reconstruídas tendem a ter diferentes frequências e orientações.

Um método para resolver este problema de estagnação é aplicar o algoritmo três vezes, cada vez com uma imagem aleatória diferente, produzir três imagens com faixas diferentes e comparar as transformadas de Fourier destas como segue: em cada ponto fazer a média dos dois valores complexos de Fourier mais próximos e descartar o terceiro. A transformada de Fourier inversa da função resultante após o processo será o próximo "output" para o algoritmo. Com mais algumas iterações obtém-se uma solução.

Outro método para sobrepor estagnações com faixas utiliza somente duas imagens com

faixas. Para cada uma destas imagens a área do suporte da imagem é tornada zero, permanecendo somente as faixas na área externa do suporte da imagem. As amplitudes das transformadas de Fourier destas faixas são usadas para determinar quais áreas no domínio de Fourier tem erro de fase associado com as faixas. Uma nova trasformada de Fourier composta é considerada sendo igual a transformada de Fourier da primeira imagem com faixas onde ela não é afetada pelos erros de fase junto com a transformada de Fourier da segunda imagem onde a primeira é afetada pelos erros de fase.

Como $f(x - x_0, y - y_0)$ tem a mesma amplitude de Fourier que f(x, y), a localização do suporte do objeto é arbitrária. Freqüentememente a imagem parcialmente reconstruída pelo algoritmo não está em perfeita concordância com as restrições de suporte. As restrições de suporte causam um corte de parte da imagem desejada, fazendo com que o algoritmo estagne. Para sair desta imagem pode-se aumentar o suporte ou então transladar o suporte ou a imagem. A quantidade de translação a ser utilizada pode ser determinada como segue: calcular a energia total da imagem de saida $g'_k(x, y)$ (isto é elevando ao quadrado e somando) sobre a área do suporte sendo utilizado e sobre este suporte transladado de um ou dois pixels em todas as direções. O suporte deverá ser transladado para a posição em que a energia foi máxima. Alternativamente, pode-se computar a "cross-correlation" de $g'_k(x, y)$ com o suporte e transladar de acordo com o pico da "cross-correlation". Isto pode ser feito ocasionalmente ou em toda iteração.

Nesta dissertação estamos propondo uma nova forma de evitar a estagnação dos algoritmos que é automática e independente do tipo de estagnação. Este novo método é sugerido pela caracterização do conjunto de soluções que descreveremos no Capítulo 4.

Capítulo 3

NOVOS MÉTODOS

Propomos neste Capítulo alguns novos métodos para o problema (P2), que também podem ser usados para (P3), que foram surgindo durante nossas pesquisas, e que, em alguns casos, tem produzido resultados melhores que os obtidos com os já existentes. Os resultados de algumas das experiências com esses algoritmos são apontados no Capítulo 7.

3.1 Função de Relaxação

É possível generalizar os algoritmos de Fienup substituindo o parâmetro β , que é constante, por uma função $\beta(k, x)$ onde k indica a iteração do algoritmo e x a coordenada do vetor $g_k(x)$. Assim, por exemplo, a iteração do algoritmo input-output híbrido passa a ser dada por

$$g_{k+1}(x) = \begin{cases} g'_k(x) & \text{se } x \in \gamma \\ g_k(x) - \beta(k, x)g'_k(x) & \text{se } x \notin \gamma \end{cases}.$$
(3.1)

Fizemos testes com o algoritmo híbrido definido desta forma considerando algumas funções $\beta(k, x)$. Em alguns casos houve divergência do algoritmo. Quando utilizamos, por exemplo, $\beta(k, x) = -g'_k(x)$ ou $\beta(k, x) = 1/k$ algumas imagens foram recuperadas com

aproximadamente o mesmo número de iterações que quando utilizado o algoritmo com o parâmetro constante (aquele parâmetro que fez com que o algoritmo convergisse mais rápido). Em outros casos o algoritmo tornou-se excessivamente lento.

Observamos também que com essa definição os algoritmos ER e OO passam a ser vistos como casos particulares do híbrido onde $\beta(k,x) = g_k(x)/g'_k(x)$ e $\beta(k,x) = g_k(x)/g'_k(x) + (\beta_o - 1)$, respectivamente. Aqui β_o indica o parâmetro constante do algoritmo output-output. Esta forma unificada de ver os algoritmos pode ser útil em futuras pesquisas.

3.2 Projeções Generalizadas Modificado

No que segue C_1 indica o conjunto dos pontos em \mathbb{R}^n (representando sinais, imagens ou objetos) que respeitam as restrições de suporte e não negatividade e C_2 indica o conjunto dos pontos de \mathbb{R}^n para os quais as amplitudes da transformada de Fourier coincidem com as amplitudes dadas. Se P_1 e P_2 são, respectivamente, as projeções sobre C_1 e C_2 , então $g_{k+1} = T_1 T_2 g_k \operatorname{com} T_i = I + \lambda_i (P_i - I) \ (i = 1, 2)$ é o algoritmo das projeções generalizadas. A função

$$J(g_k) = ||P_1g_k - g_k|| + ||P_2g_k - g_k||$$
(3.2)

é não negativa e $J(g_k) = 0$ se e somente se $g_k \in C_1 \cap C_2$. Levi e Stark [29] provaram que $J(g_{k+1}) \leq J(T_2g_k) \leq J(g_k)$ para λ_1 e λ_2 satisfazendo

$$0 \le \lambda_i \le \frac{A_i^2 + A_i}{A_i^2 + A_i - \frac{1}{2}(A_i + B_i)} = a_i$$

onde

$$A_{1} = \frac{\|P_{1}T_{2}g_{k} - T_{2}g_{k}\|}{\|P_{2}T_{2}g_{k} - T_{2}g_{k}\|},$$
$$A_{2} = \frac{\|P_{2}g_{k} - g_{k}\|}{\|P_{1}g_{k} - g_{k}\|},$$

$$B_1 = \frac{\langle P_2 T_2 g_k - T_2 g_k, P_1 T_2 g_k - T_2 g_k \rangle}{\|P_2 T_2 g_k - T_2 g_k\|^2}$$

$$B_2 = rac{\langle P_1 g_k - g_k, P_2 g_k - g_k
angle}{\|P_1 g_k - g_k\|^2}$$

Como o valor 1 está incluido no domínio dos λ_i o algoritmo ER de Fienup, onde $g_{k+1} = P_1 P_2 g_k$, tem a propriedade da redução de distância. Na realidade, neste caso, $J(g_k)$ é a raíz quadrada da distância $d^2(g_k, C_2)$ que Fienup mostrou ser não crescente com as iterações do algoritmo.

Para $\lambda_2 = 0$ $(T_1 T_2 g_k = T_1 P_2 g_k)$ obtém-se também o algoritmo OO de Fienup.

Levi e Stark sugerem dois tipos de otimização para encontrar λ_1 e λ_2 ideais:

a) Otimização por passo: encontra $\lambda_{2,ot}$ e utilizando T_2g_k encontra $\lambda_{1,ot}$.

b) Otimização por ciclo: encontra $\lambda_{1,ot}$ e $\lambda_{2,ot}$ simultaneamente.

Eles observam que a otimização por ciclo é melhor ou pelo menos igual a otimização por passo pois procura (λ_1, λ_2) num retângulo enquanto que na última a busca é sobre uma semi-reta e prova que na otimização por passo $\lambda_{i,ot} \ge 1$ (se P_1 é linear $\lambda_{1,ot} = 1$). Em suas experiências escolhem a otimização por ciclo, mas fixam $\lambda_{1,ot} = 1$ para evitar uma busca bidimensional.

Observando que no algoritmo das projeções generalizadas a relaxação da projeção P_1 faz com que as estimativas sejam não nulas fora do suporte incluímos um novo passo no algoritmo fazendo $g_{k+1} = P_1T_1T_2g_k$. Com este passo o algoritmo coincide com o utilizado por Levi e Stark em dois casos: a) quando $\lambda_1 = 1$ ou b) quando P_1 é linear (se C_1 inclue as restrições de não negatividade é claro que P_1 não é linear).

O algoritmo utilizado por Stark em suas experiências ($\lambda_1 = 1$) é um caso particular deste proposto. Ainda mais, para parâmetros λ_1 e λ_2 dentro dos mesmos limites impostos por Levi e Stark provamos que a função J permanece não crescente com as iterações do novo algoritmo. Isto pode ser verificado da seguinte forma: Utilizando a desigualdade triangular na definição de $J(T_1T_2g_k)$ obtém-se:

$$J(T_1 T_2 g_k) \ge \|P_2 T_1 T_2 g_k - P_1 T_1 T_2 g_k\|.$$
(3.3)

Como $J(P_1T_1T_2g_k) = ||P_2P_1T_1T_2g_k - P_1T_1T_2g_k|| + |P_1P_1T_1T_2g_k - P_1T_1T_2g_k|| = ||P_2P_1T_1T_2g_k - P_1T_1T_2g_k||$, e $P_2P_1T_1T_2g_k$ é a projeção de $P_1T_1T_2g_k$ sobre C_2 e $P_2T_1T_2g_k \in C_2$, tem-se

$$J(P_1T_1T_2g_k) \le ||P_2T_1T_2g_k - P_1T_1T_2g_k||.$$
(3.4)

Das equações (3.3) e (3.4) e utilizando a demonstração feita por Levi e Stark segue que:

$$J(g_k) \ge J(T_1 T_2 g_k) \ge J(P_1 T_1 T_2 g_k) = J(g_{k+1})$$
(3.5)

para qualquer estimativa g_k . Esta última equação mostra também que a cada passo (com os mesmos parâmetros), $J(P_1T_1T_2g_k) \leq J(T_1T_2g_k)$.

Capítulo 4

CARACTERIZAÇÃO DO CONJUNTO DOS OBJETOS COM AMPLITUDE DA TRANSFORMADA DE FOURIER DADA.

Neste capítulo apresentamos a caracterização do conjunto das funções discretas de uma ou duas variáveis não negativas com a amplitude da TFD dada. Apesar de que nossa ênfase ser no caso de imagens, ou seja, duas dimensões, faremos a caracterização em duas partes, uma quando o objeto considerado é um sinal unidimensional e outra quando tal objeto é uma imagem. Os cálculos de cada caso são um pouco diferentes mas os resultados são aplicáveis para ambos.

4.1 Sinais Unidimensionais

Seja $X = (x_0, x_1, ..., x_{n-1})^t \in \mathbb{R}^n$ um vetor representando a discretização de um sinal. Como a transformada rápida de Fourier trabalha com n igual a uma potência de 2, daremos maior ênfase ao caso em que n é par. Consideremos **n par**.

Seja $\lambda = (\lambda_0, \lambda_1, ..., \lambda_{n-1})^t$, com $\lambda_j = \lambda_{n-j}$, o vetor das amplitudes da transformada de Fourier dos vetores pertencentes ao conjunto a ser caracterizado. Assim estaremos investigando o conjunto

$$H = \{ Y \in \mathbb{R}^n \mid | \hat{Y}_j \mid = \lambda_j, \quad j = 0, ..., n - 1 \}$$

onde \hat{Y} representa a transformada de Fourier discreta de Y.

A transformada de Fourier discreta de um vetor $X \in H$ é dada por

$$\hat{X} = (\lambda_0 \varepsilon_0, \lambda_1 \varphi_1, ..., \lambda_{\frac{n}{2}-1} \varphi_{\frac{n}{2}-1}, \lambda_{\frac{n}{2}} \varepsilon_{\frac{n}{2}}, \lambda_{\frac{n}{2}-1} \overline{\varphi}_{\frac{n}{2}-1}, ..., \lambda_1 \overline{\varphi}_1)^t$$

com $\varepsilon_0 = \pm 1$, $\varepsilon_{\frac{n}{2}} = \pm 1$ e $\varphi_j = \cos \gamma_j + i \sin \gamma_j$. Vamos definir os seguintes vetores de $\boldsymbol{\mathcal{C}}^n$: $\hat{v}^0 = (\varepsilon_0, 0, ..., 0)^t$,

$$\hat{v}^{\frac{n}{2}} = (v_0^{\frac{n}{2}}, \dots, v_{n-1}^{\frac{n}{2}}) \operatorname{com} \hat{v}_j^{\frac{n}{2}} = \begin{cases} \varepsilon_{\frac{n}{2}} & \text{se } j = \frac{n}{2} \\ 0 & \text{em outros casos} \end{cases}$$

$$\begin{aligned} & \text{para } k = 1, \dots \frac{n}{2} - 1; \\ & \hat{u}^{k_1} = \begin{pmatrix} u_0^{k_1}, \dots, u_{n-1}^{k_1} \end{pmatrix} \quad \text{e} \quad \hat{u}^{k_2} = (u_0^{k_2}, \dots, u_{n-1}^{k_2}) \text{ onde} \\ & \hat{u}_j^{k_1} = \begin{cases} 1 \quad \text{se } j = k \quad ou \quad j = n+1-k \\ 0 \quad \text{em outros casos} \end{cases} \quad \text{e} \quad \hat{u}_j^{k_2} = \begin{cases} i \quad \text{se } j = k \\ -i \quad \text{se } j = n+1-k \\ 0 \quad \text{em outros casos} \end{cases} \end{aligned}$$

Assim

$$\hat{X} = \lambda_0 \hat{v}^0 + \lambda_{\frac{n}{2}} \hat{v}^{\frac{n}{2}} + \sum_{k=1}^{\frac{n}{2}-1} \lambda_k (\cos \gamma_k \hat{u}^{k_1} + \sin \gamma_k \hat{u}^{k_2}.)$$
(4.1)

Calculando a transformada de Fourier inversa tem-se

$$X = \lambda_0 v^0 + \lambda_{\frac{n}{2}} v^{\frac{n}{2}} + \sum_{k=1}^{\frac{n}{2}-1} \lambda_k (\cos \gamma_k u^{k1} + \sin \gamma_k u^{k2})$$
(4.2)

onde $v^0 = \varepsilon_0(1, 1, ...1)^t$, $v^{\frac{n}{2}} = \varepsilon_{\frac{n}{2}}(1, -1, 1, ..., -1)^t$, $u^{k_1} \in u^{k_2}$ $(k=1, ..., \frac{n}{2})$ são, respectivamente, as TF inversas de \hat{v}^0 , $\hat{v}^{\frac{n}{2}}$, $\hat{u}^{k_1} \in \hat{u}^{k_2}$ $(k=1, ..., \frac{n}{2})$.

A equação (4.2) representa um conjunto de quatro toróides em \mathbb{R}^n (produto cartesiano de círculos) com centros em

$$\lambda_0 v^0 + \lambda_{\frac{n}{2}} v^{\frac{n}{2}} = \lambda_0 \varepsilon_0 (1, 1, ...1)^t + \lambda_{\frac{n}{2}} \varepsilon_{\frac{n}{2}} (1, -1, 1, ..., -1)^t$$

(dependendo dos sinais de ε_0 e $\varepsilon_{\frac{n}{2}}$ que determinam v^0 e $v^{\frac{n}{2}}$). Como cada um dos toróides pertence a um sub-espaço afim de dimenção n-2, em \mathbb{R}^n , e os vetores envolvidos em (4.2) são mutuamente ortogonais, não há intersecção entre os toróides. Os valores de λ_0 e $\lambda_{\frac{n}{2}}$ determinam a proximidade entre os toróides (quanto menores forem os valores mais próximos estarão os toróides), enquanto que os valores de λ_j para $j \neq 0$ e $j \neq \frac{n}{2}$ determinam o comprimento dos eixos dos toróides e a dimensão dos toróides. A dimensão dos toróides depende destes valores no sentido que quanto mais dados nulos ($\lambda_k = 0$) menor será a dimensão dos toróides. Por exemplo, se os dados são referentes a um sinal constante ($\lambda_k = 0$ para $k \neq 0$) então os toróides se resumem a dois pontos localizados nos extremos dos vetores $\lambda_0(1,1,...1)^t \in \lambda_0(-1,-1,...-1)^t$. Se além da amplitude dada, sabe-se que o sinal X é não negativo, tem-se $\sum_{j=0}^{n-1} x_j \ge 0$ e, portanto, tal sinal pode pertencer a somente dois dos toróides citados (os que tem centro determinado por $\varepsilon_0 = 1$). Para o caso em que o sinal é constante há um único ponto (exatamente o sinal) com as amplitudes da TF dadas e portanto o problema tem solução única e, particularmente, as translações do sinal coincidem com o mesmo. Se no toróide existir um sinal não negativo X então

$$\lambda_{\frac{n}{2}} = |\sum_{j=0}^{n-1} (-1)^j x_j| \le \sum_{j=0}^{n-1} |x_j| = \lambda_0,$$

e portanto o centro destes toróide possui todas as coordenadas não negativas.

Sabe-se que se $(\hat{x}_u)_{u=0}^{n-1}$ é a TF discreta do sinal $(x_j)_{j=0}^{n-1}$ então $(\hat{x}_u.e^{\frac{-2\pi i k_u}{n}})_{u=0}^{n-1}$ é a transformada de Fourier discreta de uma translação do sinal, $(x_{j-k})_{j=0}^{n-1}$. Logo, a TFD das translações do sinal, com amplitudes da TFD dada, é

$$(\lambda_0\varepsilon_0,\lambda_1e^{\frac{-2\pi ik}{n}}\varphi_1,...,\lambda_{\frac{n}{2}-1}e^{\frac{(n-2)\pi ik}{n}}\varphi_{\frac{n}{2}-1},\lambda_{\frac{n}{2}}e^{-\pi ik}\varepsilon_{\frac{n}{2}},...,\lambda_1e^{\frac{2\pi ik}{n}}\overline{\varphi}_1)^t.$$

Como o toróide ao qual pertence a translação depende das coordenadas $\lambda_0 \varepsilon_0 e \lambda_{\frac{n}{2}} e^{-\pi i k} \varepsilon_{\frac{n}{2}}$, translações com k par estão no mesmo toróide do sinal original e translações com k ímpar estão no outro toróide em que ε_0 é o mesmo que no toróide do sinal.

Uma conclusão importante da análise anterior é que quando se quer recuperar um sinal não negativo existe um representante deste (uma translação) em cada um dos toróides onde $\varepsilon_0 = 1$.

Do ponto de vista prático, se $\lambda_0 e \lambda_{\frac{n}{2}}$ são pequenos então um sinal com amplitudes da TFD fixas pode ser visto em todos os toróides embora representado por funções diferentes. Isso pode ser verificado fácilmente observando-se que se um sinal satisfaz a equação (4.2) a simples mudança de sinal de ε_0 ou $\varepsilon_{\frac{n}{2}}$ (mudança de toróide) mostra que existe um sinal do tipo $X_2 = X_1 \pm \lambda_0 v_0 \pm \lambda_{\frac{n}{2}} v_{\frac{n}{2}}$ no outro toróide. Dependendo de qual mudança foi feita X_2 é uma suavização de X_1 ou é X_1 com um contraste maior.

Se **n** é ímpar a transformada de um vetor $X \in H$ tem a forma

$$\bar{X} = (\lambda_0 \varepsilon_0, \lambda_1 \varphi_1, ..., \lambda_{\frac{n-1}{2}} \varphi_{\frac{n-1}{2}}, \lambda_{\frac{n-1}{2}} \overline{\varphi}_{\frac{n-1}{2}}, ..., \lambda_1 \overline{\varphi}_1)^t.$$

Com um desenvolvimento análogo ao caso em que n é par conclui-se que H é um conjunto de dois toróides com centros em $\lambda_0 \varepsilon_0$ dependendo do sinal de ε_0 . Estes toróides pertencem a sub-espaços de dimensão n-1 de \mathbb{R}^n .

4.2 Imagens

Estaremos representando uma imagem por uma matriz real de ordem n (n par) com seus elementos distribuidos como na seguinte matriz:

$$X = \begin{bmatrix} x_{n-10} & x_{n-11} & \dots & x_{n-1n-1} \\ \vdots & & & \\ x_{10} & x_{11} & \dots & x_{1n-1} \\ x_{00} & x_{01} & \dots & x_{0n-1} \end{bmatrix}$$

Para n par, como a matriz é real, sua TFD é dada por

$$\hat{X} = \begin{bmatrix}
\bar{x}_{10} & \bar{x}_{1n-1} & \dots & \bar{x}_{1\frac{n}{2}} & \bar{x}_{1\frac{n}{2}-1} & \dots & \bar{x}_{11} \\
\vdots & & & & \\
\bar{x}_{\frac{n}{2}-10} & \bar{x}_{\frac{n}{2}-1n-1} & \dots & \bar{x}_{\frac{n}{2}-1\frac{n}{2}} & \bar{x}_{\frac{n}{2}-1\frac{n}{2}} & \dots & \bar{x}_{\frac{n}{2}-11} \\
\hat{x}_{\frac{n}{2}0} & \hat{x}_{\frac{n}{2}1} & \dots & \hat{x}_{\frac{n}{2}\frac{n}{2}} & \bar{x}_{\frac{n}{2}\frac{n}{2}-1} & \dots & \bar{x}_{\frac{n}{2}-11} \\
\vdots & & & & \\
\hat{x}_{10} & \hat{x}_{11} & \dots & \hat{x}_{1\frac{n}{2}} & \hat{x}_{1\frac{n}{2}+1} & \dots & \hat{x}_{1n-1} \\
\hat{x}_{00} & \hat{x}_{01} & \dots & \hat{x}_{0\frac{n}{2}} & \hat{x}_{0\frac{n}{2}-1} & \dots & \hat{x}_{01}
\end{bmatrix}$$

cujos elementos são números complexos e, em especial, \hat{x}_{00} , $\hat{x}_{0,\frac{n}{2}}$, $\hat{x}_{\frac{n}{2}0}$ e $\hat{x}_{\frac{n}{2}\frac{n}{2}}$ são reais.

Assim a matriz dos dados (amplitudes da TFD) tem a seguinte forma:

$$D = \begin{bmatrix} \lambda_{10} & \lambda_{1n-1} & \dots & \lambda_{1\frac{n}{2}} & \lambda_{1\frac{n}{2}-1} & \dots & \lambda_{11} \\ \vdots & & & & \\ \lambda_{\frac{n}{2}-10} & \lambda_{\frac{n}{2}-1n-1} & \dots & \lambda_{\frac{n}{2}-1\frac{n}{2}} & \lambda_{\frac{n}{2}\frac{n}{2}-1} & \dots & \lambda_{\frac{n}{2}1} \\ \lambda_{\frac{n}{2}0} & \lambda_{\frac{n}{2}1} & \dots & \lambda_{\frac{n}{2}\frac{n}{2}} & \lambda_{\frac{n}{2}\frac{n}{2}-1} & \dots & \lambda_{\frac{n}{2}1} \\ \vdots & & & & \\ \lambda_{10} & \lambda_{11} & \dots & \lambda_{1\frac{n}{2}} & \lambda_{1\frac{n}{2}+1} & \dots & \lambda_{1n-1} \\ \lambda_{00} & \lambda_{01} & \dots & \lambda_{0\frac{n}{2}} & \lambda_{0\frac{n}{2}-1} & \dots & \lambda_{01} \end{bmatrix}$$

Assim a TFD de uma matriz

$$X \in H = \{ Y \in \mathbb{R}^{n^2} \mid | \hat{Y}_{mj} \mid = \lambda_{mj} \quad i, j = 0, ..., n-1 \}$$

é da forma

$$\hat{X} = \begin{bmatrix} \lambda_{10}\overline{\varphi}_{10}, & \lambda_{1n-1}\overline{\varphi}_{1n-1} & \dots & \lambda_{1\frac{n}{2}}\overline{\varphi}_{1\frac{n}{2}} & \lambda_{1\frac{n}{2}-1}\overline{\varphi}_{1\frac{n}{2}-1} & \dots & \lambda_{11}\overline{\varphi}_{11} \\ \vdots & & & \\ \lambda_{\frac{n}{2}-10}\overline{\varphi}_{\frac{n}{2}-10} & \lambda_{\frac{n}{2}-1n-1}\overline{\varphi}_{\frac{n}{2}-1n-1} & \dots & \lambda_{\frac{n}{2}-1\frac{n}{2}}\overline{\varphi}_{\frac{n}{2}-1\frac{n}{2}} & \lambda_{\frac{n}{2}-1\frac{n}{2}-1}\overline{\varphi}_{\frac{n}{2}-1\frac{n}{2}-1} & \dots & \lambda_{\frac{n}{2}-11}\overline{\varphi}_{\frac{n}{2}-11} \\ \lambda_{\frac{n}{2}0}\varepsilon_{\frac{n}{2}0} & \lambda_{\frac{n}{2}1}\varphi_{\frac{n}{2}1} & \dots & \lambda_{\frac{n}{2}\frac{n}{2}}\varepsilon_{\frac{n}{2}\frac{n}{2}} & \lambda_{\frac{n}{2}\frac{n}{2}-1}\overline{\varphi}_{\frac{n}{2}\frac{n}{2}-1} & \dots & \lambda_{\frac{n}{2}1}\overline{\varphi}_{\frac{n}{2}-1} \\ \vdots & & & \\ \lambda_{10}\varphi_{10} & \lambda_{11}\varphi_{11} & \dots & \lambda_{1\frac{n}{2}}\varphi_{1\frac{n}{2}} & \lambda_{1\frac{n}{2}+1}\varphi_{1\frac{n}{2}+1} & \dots & \lambda_{1n-1}\varphi_{1n-1} \\ \lambda_{00}\varepsilon_{00} & \lambda_{01}\varphi_{01} & \dots & \lambda_{0\frac{n}{2}}\varepsilon_{0\frac{n}{2}} & \lambda_{0\frac{n}{2}-1}\overline{\varphi}_{0\frac{n}{2}-1} & \dots & \lambda_{01}\overline{\varphi}_{01} \end{bmatrix}$$

onde $\varepsilon_{00} = \pm 1$, $\varepsilon_{0\frac{n}{2}} = \pm 1$, $\varepsilon_{\frac{n}{2}0} = \pm 1$, $\varepsilon_{\frac{n}{2}\frac{n}{2}} = \pm 1$ e $\varphi_{mj} = \cos \gamma_{mj} + i \mathrm{sen} \gamma_{mj}$.

Consideremos as seguintes matrizes de \mathbb{C}^n (com os elementos distribuidos da mesma forma que nas anteriores):

,

,

,

$$\hat{A}^{00} = (\hat{a}^{00}_{mj}) \operatorname{com} \hat{a}^{00}_{mj} = \begin{cases} \varepsilon_{00} & \text{se } m = 0 & \text{e} \quad j = 0\\ 0 & \text{em outros casos} \end{cases}$$

$$\hat{A}^{0\frac{n}{2}} = (\hat{a}_{mj}^{0\frac{n}{2}}) \operatorname{com} \hat{a}_{mj}^{0\frac{n}{2}} = \begin{cases} \varepsilon_{0\frac{n}{2}} & \text{se } m = 0 & \text{e} \quad j = \frac{n}{2} \\ 0 & \text{em outros casos} \end{cases}$$

$$\hat{A}^{\frac{n}{2}0} = (\hat{a}_{mj}^{\frac{n}{2}0}) \operatorname{com} \hat{a}_{mj}^{\frac{n}{2}0} = \begin{cases} \varepsilon_{\frac{n}{2}0} & \text{se } m = \frac{n}{2} & \text{e} \quad j = 0\\ 0 & \text{em outros casos} \end{cases}$$

 \mathbf{e}

$$\hat{A}^{\frac{n}{2}\frac{n}{2}} = (\hat{a}_{mj}^{\frac{n}{2}\frac{n}{2}}) \operatorname{com} \hat{a}_{mj}^{\frac{n}{2}\frac{n}{2}} = \begin{cases} \varepsilon_{\frac{n}{2}\frac{n}{2}} & \text{se } m = \frac{n}{2} & \text{e} \quad j = \frac{n}{2} \\ 0 & \text{em outros casos} \end{cases}.$$

Para
$$k, l = 1, ..., \frac{n}{2} - 1$$
:
 $\hat{B}^{0k_1} = (\hat{b}_{mj}^{0k_1}) \operatorname{com} \hat{b}_{mj}^{0k_1} = \begin{cases} 1 & \text{se } m = 0 & \text{e} \quad (j = k & \text{ou} \quad j = n - k) \\ 0 & \text{em outros casos} \end{cases}$,

$$\hat{B}^{0k2} = (\hat{b}_{mj}^{0k2}) \operatorname{com} \hat{b}_{mj}^{0k2} = \begin{cases} i & \text{se } m = 0 & \text{e} \quad j = k \\ -i & \text{se } m = 0 & \text{e} \quad j = n - k \\ 0 & \text{em outros casos} \end{cases},$$

$$\hat{B}^{l01} = (\hat{b}_{mj}^{l01}) \operatorname{com} \hat{b}_{mj}^{l01} = \begin{cases} 1 & \text{se } j = 0 & \text{e} \quad (m = l \quad \text{ou} \quad m = n - l) \\ 0 & \text{em outros casos} \end{cases},$$

e

$$\hat{B}^{l02} = (\hat{b}_{mj}^{l02}) \operatorname{com} \hat{b}_{mj}^{l02} = \begin{cases} i & \text{se } j = 0 & \text{e} & m = l \\ -i & \text{se } j = 0 & \text{e} & m = n - l \\ 0 & \text{em outros casos} \end{cases}$$

Para $l = 1, ..., \frac{n}{2} - 1$ e k = 1, ..., n - 1:

$$\hat{B}^{lk_1} = (\hat{b}_{mj}^{lk_1})$$

$$\operatorname{com} \, \hat{b}_{mj}^{lk_1} = \begin{cases} 1 & \operatorname{se} \ (m = l \quad e \quad j = k) & \operatorname{u} \ (m = n - l \quad e \quad j = n - k) \\ 0 & \operatorname{em} \ \operatorname{outros} \ \operatorname{casos} \end{cases},$$

e

$$\hat{B}^{lk2} = (\hat{b}_{mj}^{lk2}) \operatorname{com} \hat{b}_{mj}^{lk2} = \begin{cases} i & \text{se } m = l & \text{e} \quad j = k \\ -i & \text{se } m = n - l & \text{e} \quad j = n - k \\ 0 & \text{em outros casos} \end{cases}$$

Para $l = \frac{n}{2}$ e $k = 1, ..., \frac{n}{2} - 1$:

$$\hat{B}^{lk1} = (\hat{b}_{mj}^{lk1})$$

 $\operatorname{com} \, \hat{b}_{mj}^{lk_1} = \left\{ \begin{array}{ll} 1 & \operatorname{se} \ (m=l \quad \mathrm{e} \quad j=k) \\ \\ 0 & \operatorname{em} \ \mathrm{outros} \ \mathrm{casos} \end{array} \right. \quad \mathrm{ou} \quad (m=n-l \quad \mathrm{e} \quad j=k) \\ \end{array} \right. ,$

e

$$\hat{B}^{lk2} = (\hat{b}_{mj}^{lk2}) \operatorname{com} \hat{b}_{mj}^{lk2} = \begin{cases} i & \text{se } m = l & \text{e} \quad j = k \\ -i & \text{se } m = n - l & \text{e} \quad j = k \\ 0 & \text{em outros casos} \end{cases}$$

Assim

$$\hat{X} = \lambda_{00}\hat{A}^{00} + \lambda_{0\frac{n}{2}}\hat{A}^{0\frac{n}{2}} + \lambda_{\frac{n}{2}0}\hat{A}^{\frac{n}{2}0} + \lambda_{\frac{n}{2}\frac{n}{2}}\hat{A}^{\frac{n}{2}\frac{n}{2}} + \sum_{k=1}^{\frac{n}{2}-1}\lambda_{0k}(\cos\gamma_{0k}\hat{B}^{0k1} + \sin\gamma_{0k}\hat{B}^{0k2}) \\
+ \sum_{l=1}^{\frac{n}{2}-1}\lambda_{l0}(\cos\gamma_{l0}\hat{B}^{l01} + \sin\gamma_{l0}\hat{B}^{l02}) + \sum_{l=1}^{\frac{n}{2}-1}\sum_{k=1}^{n-1}\lambda_{lk}(\cos\gamma_{lk}\hat{B}^{lk1} + \sin\gamma_{lk}\hat{B}^{lk2}) \\
+ \sum_{k=1}^{\frac{n}{2}-1}\lambda_{\frac{n}{2}k}(\cos\gamma_{\frac{n}{2}k}\hat{B}^{\frac{n}{2}k1} + sen\gamma_{\frac{n}{2}k}\hat{B}^{\frac{n}{2}k2}). \quad (4.3)$$

Calculando a TFD inversa,

$$X = \lambda_{00}A^{00} + \lambda_{0\frac{n}{2}}A^{0\frac{n}{2}} + \lambda_{\frac{n}{2}0}A^{\frac{n}{2}0} + \lambda_{\frac{n}{2}\frac{n}{2}}A^{\frac{n}{2}\frac{n}{2}} + \sum_{k=1}^{\frac{n}{2}-1}\lambda_{0k}(\cos\gamma_{0k}B^{0k1} + \sin\gamma_{0k}B^{0k2}) + \sum_{l=1}^{\frac{n}{2}-1}\lambda_{l0}(\cos\gamma_{l0}B^{l01} + \sin\gamma_{l0}B^{l02}) + \sum_{l=1}^{\frac{n}{2}-1}\sum_{k=1}^{n-1}\lambda_{lk}(\cos\gamma_{lk}B^{lk1} + \sin\gamma_{lk}B^{lk2})$$

$$+\sum_{k=1}^{\frac{n}{2}-1} \lambda_{\frac{n}{2}k} (\cos \gamma_{\frac{n}{2}k} B^{\frac{n}{2}k1} + \sin \gamma_{\frac{n}{2}k} B^{\frac{n}{2}k2}) \quad (4.4)$$

٦

onde A^{00} , $A^{0\frac{n}{2}}$, $A^{\frac{n}{2}0}$, $A^{\frac{n}{2}\frac{n}{2}}$, B^{lk_1} e B^{lk_2} são, respectivamente, as TF inversas de \hat{A}^{00} , $\hat{A}^{0\frac{n}{2}}$, $\hat{A}^{\frac{n}{2}0}$, $\hat{A}^{\frac{n}{2}\frac{n}{2}}$, \hat{B}^{lk_1} e \hat{B}^{lk_2} . Esta equação representa um conjunto de dezesseis toróides com centros em

$$\lambda_{00}A^{00} + \lambda_{0\frac{n}{2}}A^{0\frac{n}{2}} + \lambda_{\frac{n}{2}0}A^{\frac{n}{2}0} + \lambda_{\frac{n}{2}\frac{n}{2}}A^{\frac{n}{2}\frac{n}{2}} = \begin{bmatrix} c0 & c1 & \dots & c0 & \dots & c0 \\ c2 & c3 & \dots & c2 & \dots & c3 \\ c0 & c1 & \dots & c0 & \dots & c0 \\ \vdots & & & & & \\ c2 & c3 & \dots & c2 & \dots & c3 \end{bmatrix}$$
(4.5)

Ē

onde

$$c_{0} = \lambda_{00}\varepsilon_{00} + \lambda_{0\frac{n}{2}}\varepsilon_{0\frac{n}{2}} - \lambda_{\frac{n}{2}0}\varepsilon_{\frac{n}{2}0} - \lambda_{\frac{n}{2}\frac{n}{2}}\varepsilon_{\frac{n}{2}\frac{n}{2}}$$

$$c_{1} = \lambda_{00}\varepsilon_{00} - \lambda_{0\frac{n}{2}}\varepsilon_{0\frac{n}{2}} - \lambda_{\frac{n}{2}0}\varepsilon_{\frac{n}{2}0} + \lambda_{\frac{n}{2}\frac{n}{2}}\varepsilon_{\frac{n}{2}\frac{n}{2}}$$

$$c_{2} = \lambda_{00}\varepsilon_{00} + \lambda_{0\frac{n}{2}}\varepsilon_{0\frac{n}{2}} + \lambda_{\frac{n}{2}0}\varepsilon_{\frac{n}{2}0} + \lambda_{\frac{n}{2}\frac{n}{2}}\varepsilon_{\frac{n}{2}\frac{n}{2}}$$

$$c_{3} = \lambda_{00}\varepsilon_{00} - \lambda_{0\frac{n}{2}}\varepsilon_{0\frac{n}{2}} + \lambda_{\frac{n}{2}0}\varepsilon_{\frac{n}{2}0} - \lambda_{\frac{n}{2}\frac{n}{2}}\varepsilon_{\frac{n}{2}\frac{n}{2}}.$$

$$(4.6)$$

Os toróides que interceptam o conjunto dos vetores com todas as coordenadas não negativas possuem um centro com coordenadas não negativas. Isto pode ser verificado, independentemente do toróide (dos sinais dos ε_{ij}), da seguinte forma:

Sejam $P = \{0 \le i \le n-1\}, I = \{0 \le i \le n-1\}$. Para qualquer solução não negativa, $(x_{mj})_{m,j=0}^{n-1}$, pertencente ao toróide consideremos $S_1 = \sum_{j \in P} \sum_{i \in P} x_{ij}, S_2 = \sum_{j \in P} \sum_{i \in I} x_{ij}, S_3 = \sum_{j \in I} \sum_{i \in I} x_{ij}$ e $S_4 = \sum_{j \in I} \sum_{i \in I} x_{ij}$. Como, $\lambda_{00}\varepsilon_{00} = \sum_{j=0}^{n-1n-1} x_{ij}, \lambda_{\frac{n}{2}0}\varepsilon_{\frac{n}{2}0} = \sum_{j=0}^{n-1} (-1)^j \sum_{i=0}^{n-1} x_{ij},$ $\lambda_{0\frac{n}{2}}\varepsilon_{0\frac{n}{2}} = \sum_{j=0}^{n-1n-1} (-1)^i x_{ij}, e \lambda_{\frac{n}{2}\frac{n}{2}}\varepsilon_{\frac{n}{2}\frac{n}{2}} = \sum_{j=0}^{n-1} (-1)^j \sum_{i=0}^{n-1} (-1)^i x_{ij}, \text{ tem-se que } c_0 = 4S_2 \ge 0,$ $c_1 = 4S_4 \ge 0, c_2 = 4S_1 \ge 0 e disc_3 = 4S_3 \ge 0.$

Não há intercecção entre os toróides pois cada um deles pertence a um sub-espaço afim de dimensão $n^2 - 4$, em \mathbb{R}^{n^2} , e as matrizes na equação (4.3) são duas a duas ortogonais.

Aqui os valores de λ_{00} , $\lambda_{0\frac{n}{2}}$, $\lambda_{\frac{n}{2}0} \in \lambda_{\frac{n}{2}\frac{n}{2}}$ determinam a proximidade entre os toróides (quanto menores forem os valores mais próximos estarão os toróides), enquanto que os valores de λ_{lk} para os outros índices $l \in k$ determinam os comprimentos dos eixos e a dimensão dos toróides. Claramente quanto mais dados nulos menor a dimensão dos toróides. Se por exemplo, os dados são referentes a uma imagem constante então $\lambda_{lk} = 0$ para $l \neq 0$ e $k \neq 0$. Neste caso os dezesseis toróides se resumem a dois pontos dados por $\lambda_{00}\varepsilon_{00}A_{1}^{00}$, onde A_{1}^{00} é a matriz constante com todos os elementos iguais a 1.

Se a imagem cuja amplitude é dada, é também não negativa, tem-se $\sum_{m=0}^{n-1} \sum_{j=0}^{n-1} x_{mj} \ge 0$ e, portanto, tal imagem pode pertencer a oito dos toróides citados (os que tem centro determinado por $\varepsilon_{00} = 1$).

Sabe-se que se $(\hat{x}_{uv})_{u,v=0}^{n-1}$ é a transformada de Fourier discreta do sinal $(x_{mj})_{m,j=0}^{n-1}$ então $(\hat{x}_{uv}.e^{\frac{-2\pi i(au+bv)}{n}})_{u,v=0}^{n-1}$ é a TFD de uma translação da imagem, $(x_{m-a,j-b})_{m,j=0}^{n-1}$. Logo a a matriz que representa a TFD das translações da imagem, com amplitudes da TFD dada, é

 $\hat{X} =$

$$= \begin{bmatrix} \lambda_{10}e^{\frac{2\pi ia}{n}}\overline{\varphi}_{10}, & \lambda_{1n-1}e^{\frac{2\pi i(a+(n-1)b)}{n}}\overline{\varphi}_{1n-1} & \dots & \lambda_{1\frac{n}{2}}e^{\frac{2\pi i(a+\frac{n}{2}b)}{n}}\overline{\varphi}_{1\frac{n}{2}} & \dots & \lambda_{11}e^{\frac{2\pi i(a+b)}{n}}\overline{\varphi}_{11} \\ \vdots & & & \\ \lambda_{\frac{n}{2}}e^{-\pi ia}\varepsilon_{\frac{n}{2}0} & \lambda_{\frac{n}{2}1}e^{\frac{-2\pi i(\frac{n}{2}a+b)}{n}}\varphi_{\frac{n}{2}1} & \dots & \lambda_{\frac{n}{2}\frac{n}{2}}e^{-\pi i(a+b)}\varepsilon_{\frac{n}{2}\frac{n}{2}} & \dots & \lambda_{\frac{n}{2}1}e^{\frac{2\pi i(\frac{n}{2}a+b)}{n}}\overline{\varphi}_{\frac{n}{2}1} \\ \vdots & & & \\ \lambda_{10}e^{\frac{-2\pi ia}{n}}\varphi_{10} & \lambda_{11}e^{\frac{-2\pi i(a+b)}{n}}\varphi_{11} & \dots & \lambda_{1\frac{n}{2}}e^{\frac{-2\pi i(a+\frac{n}{2}b)}{n}}\varphi_{1\frac{n}{2}} & \dots & \lambda_{1n-1}e^{\frac{-2\pi i(a+(n-1)b)}{n}}\varphi_{1n-1} \\ \lambda_{00}\varepsilon_{00} & \lambda_{01}e^{\frac{-2\pi ib}{n}}\varphi_{01} & \dots & \lambda_{0\frac{n}{2}}e^{-\pi ib}\varepsilon_{0\frac{n}{2}} & \dots & \lambda_{01}e^{\frac{2\pi ib}{n}}\overline{\varphi}_{01} \end{bmatrix}$$

O toróide ao qual pertence uma translação depende do sinal dos elementos $\lambda_{00}\varepsilon_{00}$, $\lambda_{0\frac{n}{2}}e^{-\pi i b}\varepsilon_{0\frac{n}{2}}, \ \lambda_{\frac{n}{2}0}e^{-\pi i a}\varepsilon_{\frac{n}{2}0}$, e $\lambda_{\frac{n}{2}\frac{n}{2}}e^{-\pi i (a+b)}\varepsilon_{\frac{n}{2}\frac{n}{2}}$ na matriz da TFD da translação. Como $\lambda_{00}\varepsilon_{00}$ é fixo para a imagem e suas translações, estas estarão em no máximo 8 toróides. Podemos dividir as translações possíveis nos seguintes casos:

1) $a \in b$ números pares;

2) a par e b ímpar;

3) a ímpar e b par;

4) $a \in b$ números ímpares.

No primeiro caso a translação estará no mesmo toróide da imagem inicial. Nos outros três casos as translações estarão, respectivamente, em outros três toróides.

Assim uma imagem não negativa estará representada por translações dela em quatro dos oito toróides onde $\varepsilon_{00} = 1$.

Consideremos uma função f(x, y) e sua TF, F(u, v). Introduzindo coordenadas polares $x = rcos\theta$, $y = rsen\theta$, $u = w \cos \phi$ e $v = w \operatorname{sen} \phi$ é verdade que a rotação de um ângulo θ_0 na função leva a uma rotação de mesmo ângulo em sua TF, isto é, a TF de $f(r, \theta + \theta_0)$ é $F(w, \phi + \theta_0)$.

Devido a forma de discretização para as imagens (para utilizar a transformada rápida de Fourier) são possíveis somente rotações de ângulos $\theta_0 = \pi$ ou $\theta_0 = \pm \frac{\pi}{2}$. Numa rotação de ângulo π , temos que $|\hat{x}^{\pi}_{uv}| = |\hat{x}_{uv}|$ para u, v = 0, ...n - 1. Particularmente $\hat{x}^{\pi}_{00}, \hat{x}^{\pi}_{0\frac{n}{2}}, \hat{x}^{\pi}_{\frac{n}{2}0}$ e $\hat{x}^{\pi}_{\frac{n}{2}\frac{n}{2}}$ são, respectivamente, iguais a $\hat{x}_{00}, \hat{x}_{0\frac{n}{2}}, \hat{x}^{\pi}_{\frac{n}{2}0}$ e $\hat{x}_{\frac{n}{2}\frac{n}{2}}$. Logo rotações de ângulo π de uma imagem pertencem ao mesmo toróide da imagem. Em geral, as amplitudes da TF não são mantidas com rotações de ângulo $\theta_0 = \pm \frac{\pi}{2}$. Mesmo em casos especiais, onde as restrições de amplitude são mantidas, a rotação pertencerá a um dos toróides que contém translações da imagem.

Conclui-se que uma imagem não negativa, com amplitude da TF dada, é representada por suas translações ou rotações em quatro dos oito toróides onde $\varepsilon_{00} = 1$.

Análogamente ao caso de sinais, na prática, se λ_{00} , $\lambda_{0\frac{n}{2}}$, $\lambda_{\frac{n}{2}0} \in \lambda_{\frac{n}{2}\frac{n}{2}}$ são pequenos então uma imagem pode ser vista em todos os toróides embora representada por funções diferentes. Isso pode ser verificado fácilmente observando-se que se uma imagem satisfaz a equação (4.4) a simples mudança de sinal de um dos ε_{lm} (mudança de toróide) mostra que existe uma imagem em outro toróide que, dependendo de qual mudança foi feita, é a original com um contraste maior ou menor. Se **n** é ímpar a transformada de um vetor $X \in H$ tem a forma

$$\hat{X} = \begin{bmatrix} \lambda_{10}\overline{\varphi}_{10} & \dots & \lambda_{1\frac{n+1}{2}}\overline{\varphi}_{1\frac{n+1}{2}} & \lambda_{1\frac{n-1}{2}-1}\overline{\varphi}_{1\frac{n-1}{2}-1} & \dots & \lambda_{11}\overline{\varphi}_{11} \\ \vdots \\ \lambda_{\frac{n-1}{2}0}\overline{\varphi}_{\frac{n-1}{2}0} & \dots & \lambda_{\frac{n-1}{2}\frac{n+1}{2}}\overline{\varphi}_{\frac{n-1}{2}\frac{n+1}{2}} & \lambda_{\frac{n-1}{2}\frac{n-1}{2}}\overline{\varphi}_{\frac{n-1}{2}\frac{n-1}{2}} & \dots & \lambda_{\frac{n-1}{2}1}\overline{\varphi}_{\frac{n-1}{2}1} \\ \lambda_{\frac{n-1}{2}0}\varphi_{\frac{n-1}{2}0} & \dots & \lambda_{\frac{n-1}{2}\frac{n-1}{2}}\varphi_{\frac{n-1}{2}\frac{n-1}{2}} & \lambda_{\frac{n-1}{2}\frac{n+1}{2}}\overline{\varphi}_{\frac{n-1}{2}\frac{n+1}{2}} & \dots & \lambda_{\frac{n-1}{2}n-1}\overline{\varphi}_{\frac{n+1}{2}n-1} \\ \vdots \\ \lambda_{10}\varphi_{10} & \dots & \lambda_{1\frac{n-1}{2}}\varphi_{1\frac{n-1}{2}} & \lambda_{1\frac{n-1}{2}}\varphi_{1\frac{n+1}{2}} & \dots & \lambda_{1n-1}\varphi_{1n-1} \\ \lambda_{00}\varepsilon_{00} & \dots & \lambda_{0\frac{n-1}{2}}\varphi_{0\frac{n-1}{2}} & \lambda_{0\frac{n-1}{2}}\overline{\varphi}_{0\frac{n-1}{2}} & \dots & \lambda_{01}\overline{\varphi}_{01} \end{bmatrix}$$

Com um desenvolvimento análogo ao caso em que n é par conclui-se que H é um conjunto de dois toróides caracterizados pelo sinal de ε_{00} . Estes toróides pertencem a sub-espaços de dimensão $n^2 - 1$ de \mathbb{R}^{n^2} .

A caracterização descrita neste capítulo é o fundamento para as demonstrações do Capítulo 5 assim como para a nova heurística para resolver o problema das estagnações sugerida no Capítulo 6.

Capítulo 5

PROPRIEDADES DE CONVERGÊNCIA

Está claro que os métodos iterativos descritos no Capítulo 3 não necessariamente convergem para uma solução do problema e, portanto, a análise das propriedades de convergência deve se restringir, no caso global, a provar que a seqüência gerada pelos mesmos é limitada também no domínio do espaço (no domínio das frequências isto é óbvio, já que estamos projetando sempre sobre um conjunto compacto [trataremos apenas o problema em dimensão finita, tal como dito anteriormente]). A prova de limitação pelo menos garante que os algoritmos geram sempre subseqüências convergentes. Este é um passo fundamental para qualquer análise posterior. Neste capítulo provamos que o algoritmos IO e IOH são limitados no domínio do espaço quando é usada a restrição de não negatividade na definição do passo do algoritmo e após este é calculada a projeção sobre a Uma continuação deste trabalho seria provar a convergência local restrição de suporte. (que conjecturamos para valores suficientemente pequenos de β , no caso do IOH e OO), e a convergência das subseqüências que correspondem aos pontos fixos dos algoritmos (uma tarefa difícil). Na última seção analisamos os pontos fixos dos algoritmos que não necessariamente são soluções.

5.1 Limitação dos Algoritmos

5.1.1 Error-Reduction

Consideremos C_1 o conjunto dos vetores que satisfazem as restrições no espaço tempo e C_2 o conjunto dos vetores cuja TF tem a magnitude dada. Se P_1 e P_2 são, respectivamente, as projeções sobre C_1 e C_2 a iteração do algoritmo ER é dada por $g_{k+1} = P_1P_2g_k$. Levi e Stark [29] definiram a função $J(g_k) = ||P_1g_k - g_k|| + ||P_2g_k - g_k||$ e provaram que, em particular, a mesma é não crescente com as iterações deste algoritmo. Na verdade Fienup [17] já havia obtido este resultado, pois para este algoritmo, $J(g_k) = ||P_2g_k - g_k||$. Como C_2 é compacto, a seqüência $\{g_k\}$ é limitada.

5.1.2 Output-Output

Considerando as definições feitas no caso do algoritmo ER e $T_1 = I + \lambda_i (P_1 - I)$ tem-se que a iteração deste algoritmo é dada por $g_{k+1} = T_1 P_2 g_k$. Portanto o algoritmo OO é o algoritmo ER com a projeção P_1 relaxada. Assim, para qualquer parâmetro β fixo, a seqüência $\{g_k\}$ é limitada.

5.1.3 Input-Output Híbrido

Demonstraremos que este algoritmo é limitado no caso particular, e muito importante, em que as restrições do espaço tempo são somente restrições de não-negatividade do sinal ou imagem. Observamos que se o suporte é conhecido e for introduzido na forma de uma projeção após o passo do algoritmo a limitação continua verdadeira. Para esta prova foi preciso, novamente, como no capítulo anterior, considerar separadamente recuperação de sinais unidimensionais e recuperação de imagens.

Recuperação de Sinais Unidimensionais

Consideremos $g_0 \in \mathbb{R}^n$ dada. O algoritmo híbrido é definido assim: Se constroem duas seqüências em \mathbb{R}^n , $\{g_i\} \in \{g'_i\}$, da seguinte forma:

$$g_i = P_H g_i$$

onde

$$H = \{ X \mid |\hat{X}| = \lambda \},\$$

 λ é o vetor das amplitudes e \hat{X} é a TFD de X, isto é $\hat{X}_k = \sum_{m=0}^{n-1} X_m exp[-j(k.m)/n]$ para k = 0, ..., n-1.

(2)

$$g_{i+1,k} = \begin{cases} g_{i,k}' & \text{se } g_{i,k}' \ge 0\\ g_{i,k} - \beta g_{i,k}' & \text{em outro caso} \end{cases}$$

com $\beta \geq 0$;

Teorema 1 O algoritmo híbrido é limitado, isto é, $\{g_i\}$ é limitada.

Lema 1 A sequência $\{g_i\}$ é não negativa se $g_0 \ge 0$.

Prova:

Seja $m \in \{1, 2, ..., n\}$ o índice que indica a coordenada de uma estimativa g_i . Vamos

provar por indução que $g_{i,m} \ge 0$ para todo $m \in \{1, 2, ..., n\}$. Por hipótese $g_{0,m} \ge 0$ para m = 1, ..., n. Para $i \ge 0$, (a) se $g'_{i,m} \ge 0$ tem-se $g_{i+1,m} = g'_{i,m} \ge 0$; (b) se $g'_{i,m} < 0$ tem-se $g_{i+1,m} = g_{i,m} - \beta g'_{i,m} \ge g_{i,m} \ge 0$.

C.Q.D.

Embora com algumas diferenças nas definições quando n é ímpar ou par as demonstrações que seguem são iguais para ambos os casos.

Para as próximas definições n é ímpar. Seja e = (1, 1, ..., 1) o vetor ortogonal aos subespaços afins de \mathbb{R}^n que contém os toróides. Consideraremos a projeção $\tilde{X} = X - \alpha_X e$, com

$$\alpha_X = \frac{\langle X, e \rangle}{\langle e, e \rangle} = \frac{\sum_{i=0}^{n-1} X_i}{n},$$

de $X \in \mathbb{R}^n$ sobre o sub-espaço ortogonal ao vetor e. Equivalentemente, no espaço de Fourier tem-se a projeção sobre o sub-espaço ortogonal ao vetor $\hat{e} = (n, 0, ..., 0)$ que é dada por

$$\tilde{\hat{X}} = \hat{X} - \alpha_X \hat{e}.$$

O ângulo entre as projeções de X e Y (\tilde{X} e $\tilde{Y}),\,\sigma(\tilde{X},\tilde{Y}),$ é dado por

$$\cos \sigma(\tilde{X}, \tilde{Y}) = \frac{\langle \tilde{X}, \tilde{Y} \rangle}{\|\tilde{X}\|_2 \|\tilde{Y}\|_2},$$

onde

$$\langle \tilde{X}, \tilde{Y} \rangle = \langle X, Y \rangle - n\alpha_X \alpha_Y = \frac{\langle \hat{X}, \hat{Y} \rangle - \hat{X}_0 \hat{Y}_0}{n^2}$$

e

$$\|\tilde{X}\|_{2}^{2} = \|X\|_{2}^{2} - n\alpha_{X}^{2} = \frac{\|\hat{X}\|_{2}^{2} - |\hat{X}_{0}|^{2}}{n^{2}}$$

Para n par, consideraremos

$$\tilde{X} = X - \alpha_X e_1 - \gamma_X e_2$$

com

$$\gamma_X = \frac{\langle X, e_2 \rangle}{\langle e_2, e_2 \rangle} = \frac{\sum_{i=0}^{n-1} (-1)^i X_i}{n} = \frac{\hat{X}_{n/2}}{n},$$

a projeção de X sobre o sub-espaço ortogonal aos vetores $e_1 = (1, 1, ..., 1)$ e $e_2 = (1, -1, ..., 1, -1)$ (que são ortogonais aos sub-espaços afins que contém os toróides). Equivalentemente, no espaço de Fourier tem-se a projeção sobre o sub-espaço ortogonal aos vetores $\hat{e}_1 = (n, 0, ..., 0)$ e $\hat{e}_2 = (0, ..., n, ...0)$ que é dada por

$$\tilde{\hat{X}} = \hat{X} - \alpha_X \hat{e}_1 - \gamma_X \hat{e}_2.$$

Neste caso

$$\langle \tilde{X}, \tilde{Y} \rangle = \langle X, Y \rangle - n\alpha_X \alpha_Y - n\gamma_X \gamma_Y = \frac{\langle \hat{X}, \hat{Y} \rangle - \hat{X}_0 \hat{Y}_0 - \hat{X}_{\frac{n}{2}} \hat{Y}_{\frac{n}{2}}}{n^2}$$

e

$$\|\tilde{X}\|_{2}^{2} = \|X\|_{2}^{2} - n\alpha_{X}^{2} - n\gamma_{Y}^{2} = \frac{\|\hat{X}\|_{2}^{2} - |\hat{X}_{0}|^{2} - |\hat{X}_{\frac{n}{2}}|^{2}}{n^{2}}.$$

Sejam

$$g_{i} = \begin{bmatrix} X_{0} \\ X_{1} \\ \vdots \\ X_{n-1} \end{bmatrix}, \quad \hat{g}_{i} = \begin{bmatrix} \mu_{0} \\ \mu_{1}e^{i\delta_{1}} \\ \mu_{2}e^{i\delta_{2}} \\ \vdots \\ \mu_{1}\overline{e^{i\delta_{1}}} \end{bmatrix}, \quad g_{i-1}^{'} = \begin{bmatrix} Y_{0} \\ Y_{1} \\ \vdots \\ Y_{n-1} \end{bmatrix}, \quad \hat{g}_{i-1}^{'} = \begin{bmatrix} \lambda_{0} \\ \lambda_{1}e^{i\gamma_{1}} \\ \lambda_{2}e^{i\gamma_{2}} \\ \vdots \\ \lambda_{1}\overline{e^{i\gamma_{1}}} \end{bmatrix}$$

onde μ_m , δ_m e γ_m representam, respectivamente, amplitudes na *i*-ésima iteração, fases na *i*-ésima iteração e fases na iteração i - 1. Os λ_m são as amplitudes corretas (dados).

Tem-se
$$\alpha_{g_i} = \frac{\mu_0}{n}, \ \alpha_{g'_{i-1}} = \frac{\lambda_0}{n}, \ \gamma_{g_i} = \frac{\mu_{\frac{n}{2}}\overline{\varepsilon}_{\frac{n}{2}}}{n} \ \mathrm{e} \ \gamma_{g'_{i-1}} = \frac{\lambda_{\frac{n}{2}}\varepsilon_{\frac{n}{2}}}{n}, \ \mathrm{onde} \ \overline{\varepsilon}_{\frac{n}{2}} = \pm 1 \ \mathrm{e} \ \varepsilon_{\frac{n}{2}} = \pm 1$$

Lema 2 Se a seqüência $\{g_i\}$ não é limitada e $g_0 \ge 0$, então existem uma subseqüência $\{g_{i_p}\}$ com $i_p \in \mathbb{N}$, $p = 1, 2, ..., i_p < i_{p+1}$, um subconjunto $\Omega \subset \{0, 1, ..., n-1\}$ e uma

constante K tais que:

1) Para $m \in \Omega$, $\begin{cases} g_{i_p,m} \to +\infty, \quad g'_{i_p-1,m} < 0 \\ g_{i_p-1,m} = g_{i_p,m} + \beta g'_{i_p-1,m} \end{cases}$ 2) Para $m \notin \Omega$ $\begin{cases} g_{i_p,m} < K \\ g_{i_p-1,m} < K \end{cases}$ 3) Para todo m:

$$|g_{i_p-1,m}| < K$$

4) Se
$$n \in par$$
, o sinal de

$$\mu_{\frac{n}{2}}\overline{\varepsilon}_{\frac{n}{2}} = \sum_{m=0}^{n-1} (-1)^m g_i(m)$$

é fixo;

Prova:

Como a seqüência $\{g_i\}$ não é limitada, $\{||g_i||_1\}$ não é limitada e pode-se escolher uma subseqüência tal que $||g_{i-1}||_1 < ||g_i||_1 \rightarrow +\infty$. Considerando m=0, se existe K_1 tal que $|g_{i,1}| < K_1$ para todo i declara-se que $0 \notin \Omega$; Se não, pode-se escolher uma subseqüência da subseqüência anterior para a qual $|g_{i,1}| \rightarrow +\infty$ e declarar $0 \in \Omega$. Fazendo isso, indutivamente, para todo $m \in \{0, 1, ..., n-1\}$ e escolhendo K suficientemente grande o resultado é uma subseqüência que satisfaz (1) e (2). Como A é compacto a condição (3) é satisfeita. Para o caso em que n é par basta escolher uma subseqüência da anterior com sinal de $\mu_{\frac{n}{2}}$ fixo.

C.Q.D.

No que segue, sempre que citarmos $g_{i,m}$ estaremos nos referindo a subseqüência do lema anterior.

Lema 3

$$\cos \sigma(\tilde{g}_{i-1}, \tilde{g}'_{i-1}) \ge \frac{\lambda}{\sqrt{[\frac{n-1}{2}]}\overline{\lambda}}$$

onde

$$\overline{\lambda} = \max_{1 \le m \le [rac{n-1}{2}]} \{\lambda_m\} \quad e \quad \underline{\lambda} = \min_{1 \le m \le [rac{n-1}{2}]} \{\lambda_m\}.$$

Prova:

Verifica-se que

$$\|\tilde{g}_{i-1}'\|_{2}^{2} = \frac{2}{n^{2}} \sum_{m=1}^{\left[\frac{n-1}{2}\right]} \lambda_{m}^{2}, \quad \|\tilde{g}_{i-1}\|_{2}^{2} = \frac{2}{n^{2}} \sum_{m=1}^{\left[\frac{n-1}{2}\right]} \mu_{m}^{2} \quad e \quad \langle \tilde{g}_{i-1}', \tilde{g}_{i-1} \rangle = \frac{2}{n^{2}} \sum_{m=1}^{\left[\frac{n-1}{2}\right]} \lambda_{m} \mu_{m}.$$

Por estimativas simples,

$$\cos \sigma(\tilde{g}_{i-1}, \tilde{g}_{i-1}') = \frac{\langle \tilde{g}_{i-1}', \tilde{g}_{i-1} \rangle}{\|\tilde{g}_{i-1}'\|_2 \|\tilde{g}_{i-1}\|_2} \ge \frac{\lambda}{\sqrt{[\frac{n-1}{2}]} \overline{\lambda}} \cdot \frac{\sum_{m=1}^{[\frac{n-1}{2}]} \mu_m}{\left[\sum_{m=1}^{[\frac{n-1}{2}]} \mu_m^2\right]^{1/2}} \ge \frac{\lambda}{\sqrt{[\frac{n-1}{2}]} \overline{\lambda}}.$$

C.Q.D.

Lema 4 Para i muito grande,

$$\cos \sigma(\tilde{g}'_{i-1}, \tilde{g}_i) < -\frac{\lambda_0}{[2\sum_{m=1}^{[\frac{n-1}{2}]} \lambda_m^2]^{1/2}}.$$

Prova:

Temos que para n ímpar,

$$\langle \tilde{g}'_{i-1}, \tilde{g}_i \rangle = \langle g'_{i-1}, g_i \rangle - \frac{\lambda_0}{n} \mu_0,$$

$$||\tilde{g}_i||_2 = \left[\sum_{m=0}^{n-1} X_m^2 - \frac{1}{n} \left(\sum_{m=0}^{n-1} X_m \right)^2 \right]^{1/2}$$

enquanto que para n par,

$$\langle \tilde{g}'_{i-1}, \tilde{g}_i \rangle = \langle g'_{i-1}, g_i \rangle - \frac{\lambda_0}{n} \mu_0 - \frac{\lambda_{\frac{n}{2}} \varepsilon_{\frac{n}{2}}}{n} \mu_{\frac{n}{2}} \overline{\varepsilon}_{\frac{n}{2}},$$

е

$$\|\tilde{g}_i\|_2 = \left[\sum_{m=0}^{n-1} X_m^2 - \frac{1}{n} \left(\sum_{m=0}^{n-1} X_m\right)^2 - \frac{1}{n} \left(\sum_{m=0}^{n-1} (-1)^m X_m\right)^2\right]^{1/2}.$$

Em ambos os casos,

$$\begin{aligned} \cos \sigma(\tilde{g}_{i-1}', \tilde{g}_{i}) &= \frac{\langle \tilde{g}_{i-1}', \tilde{g}_{i} \rangle}{||\tilde{g}_{i-1}'||_{2} ||\tilde{g}_{i}||_{2}} \leq \frac{\langle g_{i-1}', g_{i} \rangle - \frac{\lambda_{0}}{n} \mu_{0}}{||\tilde{g}_{i-1}'||_{2} ||\tilde{g}_{i}||_{2}} \\ &= \frac{\langle g_{i-1}', g_{i} \rangle}{||\tilde{g}_{i-1}'||_{2} ||\tilde{g}_{i}||_{2}} - \frac{\frac{\lambda_{0}}{n} \mu_{0}}{||\tilde{g}_{i-1}'||_{2} ||\tilde{g}_{i}||_{2}} \\ &\leq \frac{\langle g_{i-1}', g_{i} \rangle}{||\tilde{g}_{i-1}'||_{2} ||\tilde{g}_{i}||_{2}} - \frac{\frac{\lambda_{0}}{n} \sum_{m=0}^{n-1} X_{m}}{||\tilde{g}_{i-1}'||_{2} \left[\sum_{m=0}^{n-1} X_{m}^{2}\right]^{1/2}} \\ &\leq \frac{\sum_{m \in \Omega} Y_{m} X_{m} + \sum_{m \notin \Omega} Y_{m} X_{m}}{||\tilde{g}_{i-1}'||_{2} \left[\sum_{m=0}^{n-1} X_{m}^{2}\right]^{1/2}} - \frac{\frac{\lambda_{0}}{n}}{\left[\frac{2}{n^{2}} \sum_{m=1}^{n-1} \lambda_{m}^{2}\right]^{1/2}}.\end{aligned}$$

Como,

$$\sum_{m \in \Omega} Y_m X_m \to -\infty, \quad e \quad \sum_{m \notin \Omega} Y_m X_m$$

é limitada, para i suficientemente grande,

$$\frac{\langle g'_{i-1}, \tilde{g}_i \rangle}{\|\tilde{g}'_{i-1}\|_2 \|\tilde{g}_i\|_2} < -\frac{\lambda_0}{\left[2\sum_{m=1}^{\lfloor \frac{n-1}{2} \rfloor} \lambda_m^2\right]^{1/2}}.$$

C.Q.D.

Lema 5

$$\lim_{i \to \infty} \frac{\|\tilde{g}_{i-1}\|_2}{\|\tilde{g}_i\|_2} = 1.$$

Prova:

Para n impar:

Podemos escrever

$$\|\tilde{g}_i\|_2^2 = \sum_{m=0}^{n-1} X_m^2 - \frac{1}{n} (\sum_{m=0}^{n-1} X_m)^2 = A + B$$

onde

$$A = \left[\sum_{m \in \Omega} X_m^2 - \frac{1}{n} (\sum_{m \in \Omega} X_m)^2\right] \quad e \quad B = \left[\sum_{m \notin \Omega} X_m^2 - \frac{1}{n} \sum_{m \notin \Omega} X_m \left(\sum_{m \notin \Omega} X_m + 2\sum_{m \in \Omega} X_m\right)\right].$$

Vamos provar que para i muito grande,

$$B \ll \frac{\sum_{m \in \Omega} X_m^2}{n} < A.$$

Pela construção da subseqüência (Lema 2) tem-se que 0 < ω =
| Ω |< n. Em primeiro lugar,

$$A = \frac{1}{n} \left\{ \omega \sum_{m \in \Omega} X_m^2 - (\sum_{m \in \Omega} X_m)^2 \right\} + \frac{n - \omega}{n} \sum_{m \in \Omega} X_m^2$$
$$\geq \frac{\sum_{m \in \Omega} X_m^2}{n}$$

pois

$$\omega \sum_{m \in \Omega} X_m^2 - (\sum_{m \in \Omega} X_m)^2 \ge 0.$$

Pela limitação dos elementos e por estimativas simples, baseadas no Lema 2,

$$B \le (n-\omega)K^2 + \frac{n-\omega}{n}K\left((n-\omega)K + 2\sum_{m\in\Omega}X_m\right) = L_0 + L_1\sum_{m\in\Omega}X_m.$$

Como, para i muito grande,

$$\sum_{m\in\Omega} X_m \ll \frac{1}{n} \sum_{m\in\Omega} X_m^2$$

vem que

$$B \ll \frac{1}{n} \sum_{m \in \Omega} X_m^2 \le A$$

Similarmente, definindo $Z_m = g_{i-1,m} = X_m + \beta Y_m$, pode-se escrever

$$\|\tilde{g}_{i-1}\|_2^2 = \sum_{m=0}^{n-1} Z_m^2 - \frac{1}{n} \left(\sum_{m=0}^{n-1} Z_m\right)^2 = A + C$$

onde

$$C = \beta \sum_{m \in \Omega} Y_m \left(\beta Y_m + 2X_m\right) + \sum_{m \notin \Omega} Z_m^2$$
$$-\frac{\beta}{n} \left(\sum_{m \in \Omega} Y_m\right) \left(\beta \sum_{m \in \Omega} Y_m + 2\sum_{m \in \Omega} X_m\right) - \frac{1}{n} \left(\sum_{m \in \Omega} Z_m\right)^2$$

Por estimativas similares às feitas no denominador,

$$C \le M_0 + M_1 \sum_{m \in \Omega} X_m \ll \frac{1}{n} \sum_{m \in \Omega} X_m^2$$

Consequentemente,

$$\lim_{i \to \infty} \frac{\|\tilde{g}_{i-1}\|_2}{\|\tilde{g}_i\|_2} = \lim_{i \to \infty} \sqrt{\frac{A+C}{A+B}} = 1.$$

Para n par:

Pela construção da subseqüência no Lema 2 tem-se 0 < $\omega = \mid \Omega \mid < n.$

Podemos escrever:

$$\|\tilde{g}_i\|_2^2 = \sum_{m=0}^{n-1} X_m^2 - \frac{1}{n} \left(\sum_{m=0}^{n-1} X_m\right)^2 - \frac{1}{n} \left(\sum_{m=0}^{n-1} (-1)^m X_m\right)^2 = A + B$$

onde

$$A = \sum_{m \in \Omega} X_m^2 - \frac{1}{n} \left(\sum_{m \in \Omega} X_m \right)^2$$

e

$$B = \left[\sum_{m \notin \Omega} X_m^2 - \frac{1}{n} \sum_{m \notin \Omega} X_m \left(\sum_{m \notin \Omega} X_m + 2 \sum_{m \in \Omega} X_m \right) \right]$$

$$-\frac{1}{n}\left(\sum_{m=0}^{n-1}(-1)^m X_m\right)^2.$$

Vamos provar que para i muito grande,

$$B \ll \frac{\sum_{m \in \Omega} X_m^2}{n} < A$$

Em primeiro lugar,

$$A = \frac{1}{n} \left\{ \omega \sum_{m \in \Omega} X_m^2 - \left(\sum_{m \in \Omega} X_m \right)^2 \right\} + \frac{n - \omega}{n} \sum_{m \in \Omega} X_m^2$$
$$\geq \frac{\sum_{m \in \Omega} X_m^2}{n}$$

pois,

$$\omega \sum_{m \in \Omega} X_m^2 - (\sum_{m \in \Omega} X_m)^2 \ge 0.$$

Pela limitação dos elementos e por estimativas simples, baseadas no Lema 2,

$$B \le (n-\omega)K^2 + \frac{2(n-\omega)}{n}K\left((n-\omega)K + 2\sum_{m\in\Omega}X_m\right) = L_0 + L_1\sum_{m\in\Omega}X_m$$

Como, para i muito grande,

$$\sum_{m\in\Omega} X_m \ll \frac{1}{n} \sum_{m\in\Omega} X_m^2$$

vem que

$$B \ll \frac{1}{n} \sum_{m \in \Omega} X_m^2 \le A.$$

Pode-se escrever

$$\|\tilde{g}_{i-1}\|_2^2 = \sum_{m=0}^{n-1} Z_m^2 - \frac{1}{n} \left(\sum_{m=0}^{n-1} Z_m\right)^2 - \frac{1}{n} \left(\sum_{m=0}^{n-1} (-1)^m Z_m\right)^2 = A + C$$

onde

$$C = \beta \sum_{m \in \Omega} Y_m \left(\beta Y_m + 2X_m \right) + \sum_{m \notin \Omega} Z_m^2$$

$$-\frac{\beta}{n}\left(\sum_{m\in\Omega}Y_m\right)\left(\beta\sum_{m\in\Omega}Y_m+2\sum_{m\in\Omega}X_m\right)-\frac{1}{n}\left(\sum_{m\in\Omega}Z_m\right)^2-\frac{1}{n}\left(\sum_{m=0}^{n-1}(-1)^mZ_m\right)^2.$$

Por estimativas similares as feitas no denominador,

$$C \leq M_0 + M_1 \sum_{m \in \Omega} X_m \ll \frac{1}{n} \sum_{m \in \Omega} X_m^2.$$

Consequentemente,

$$\lim_{i \to \infty} \frac{\|\tilde{g}_{i-1}\|_2}{\|\tilde{g}_i\|_2} = \lim_{i \to \infty} \sqrt{\frac{A+C}{A+B}} = 1.$$

Lema 6

$$\lim_{i \to \infty} \cos \sigma(\tilde{g}_{i-1}, \tilde{g}_i) = 1.$$

Prova:

$$\lim_{i \to \infty} \frac{\langle \tilde{g}_{i-1}, \tilde{g}_i \rangle}{\|\tilde{g}_{i-1}\|_2 \|\tilde{g}_i\|_2}$$

$$\geq \lim_{i \to \infty} \frac{\|\tilde{g}_i\|_2^2 - 3K\beta \sum_{m=0}^{n-1} X_m}{\|\tilde{g}_i\|_2^2} \cdot \lim_{i \to \infty} \frac{\|\tilde{g}_i\|_2}{\|\tilde{g}_{i-1}\|_2}$$

$$= \lim_{i \to \infty} \frac{\|\tilde{g}_i\|_2^2 - 3K\beta \sum_{m=0}^{n-1} X_m}{\|\tilde{g}_i\|_2^2}.$$

Logo,

$$\lim_{i \to \infty} \cos \sigma(\tilde{g}_{i-1}, \tilde{g}_i) = 1$$

C.Q.D.

PROVA DO TEOREMA 1

Supondo que a seqüência gerada pelo algoritmo não é limitada: Pelo Lema 3,

$$\sigma_1^i = \sigma(\tilde{g}_{i-1}, \tilde{g}'_{i-1}) \le \frac{\pi}{2};$$

Pelo Lema 4,

$$\sigma_{2}^{i} = \cos \sigma(\tilde{g}_{i-1}^{'}, \tilde{g}_{i}) < -\frac{\lambda_{0}}{[2\sum_{m=1}^{[n/2]} \lambda_{m}^{2}]^{1/2}}$$

e como o termo a direita desta desigualdade não depende da iteração,

$$\sigma_2^i \geq \frac{\pi}{2} + \epsilon$$

onde ϵ é uma constante estritamente positiva; Pelo Lema 6,

$$\lim_{i \to \infty} \sigma_3^i = \lim_{i \to \infty} \sigma(\tilde{g}_{i-1}, \tilde{g}_i) = 0$$

Então, para $i \gg 0$, $\sigma_1^i + \sigma_3^i < \sigma_2^i$ que é uma contradição.

C.Q.D.

Recuperação de Imagens

O algoritmo híbrido para recuperar uma imagem não negativa, $f \in \mathbb{R}^{n^2}$, dadas as amplitudes de sua TF é definido da mesma forma que para recuperação de sinais. Sejam $\{g_i\} \in \{g'_i\}$ as duas seqüências em \mathbb{R}^{n^2} geradas pelo algoritmo. Verifica-se trivialmente que o Lema 1 é válido aqui.

Faremos as demonstrações necessárias para verificar a validade do Teorema 1 neste caso. Para isso consideraremos n par (que é o caso de maior interesse).

Consideremos as seguintes matrizes:

 $E^{00}=(e_{mj}^{00})$ com $e_{mj}^{00}=1$ para quaisquer m e j.

$$E^{\frac{n}{2}} = (e_{mj}^{\frac{n}{2}}) \operatorname{com} e_{mj}^{\frac{n}{2}} = \begin{cases} 1 & \text{se j \acute{e} par} \\ -1 & \text{em outro caso} \end{cases},$$
$$E^{\frac{n}{2}0} = (e_{mj}^{\frac{n}{2}0}) \operatorname{com} e_{mj}^{\frac{n}{2}0} = \begin{cases} 1 & \text{se m \acute{e} par} \\ -1 & \text{em outro caso} \end{cases},$$
$$e & E^{\frac{n}{2}\frac{n}{2}} = (e_{mj}^{\frac{n}{2}\frac{n}{2}}) \operatorname{com} e_{mj}^{\frac{n}{2}\frac{n}{2}} = \begin{cases} 1 & \text{se m+j \acute{e} par} \\ -1 & \text{em outro caso} \end{cases}$$

A projeção de $X \in I\!\!R^{n^2}$ sobre o sub-espaço ortogonal a $E^{00}, E^{0\frac{n}{2}}, E^{\frac{n}{2}0}$ e $E^{\frac{n}{2}\frac{n}{2}}$ é dada por

$$\tilde{X} = X - \alpha_X^{00} E^{00} - \alpha_X^{0\frac{n}{2}} E^{0\frac{n}{2}} - \alpha_X^{\frac{n}{2}0} E^{\frac{n}{2}0} - \alpha_X^{\frac{n}{2}\frac{n}{2}} E^{\frac{n}{2}\frac{n}{2}}$$

onde

$$\alpha_X^{00} = \frac{\langle X, E^{00} \rangle}{\langle E_{00}, E_{00} \rangle} = \frac{1}{n^2} \sum_{m=0}^{n-1} \sum_{j=0}^{n-1} X_{mj} = \frac{\hat{X}_{00}}{n^2},$$

$$\alpha_X^{0\frac{n}{2}} = \frac{\langle X, E^{0\frac{n}{2}} \rangle}{\langle E_{0\frac{n}{2}}, E_{0\frac{n}{2}} \rangle} = \frac{1}{n^2} \sum_{m=0}^{n-1} \sum_{j=0}^{n-1} (-1)^j X_{mj} = \frac{\hat{X}_{0\frac{n}{2}}}{n^2},$$

$$\alpha_X^{\frac{n}{2}0} = \frac{\langle X, E^{\frac{n}{2}0} \rangle}{\langle E_{\frac{n}{2}0}, E_{\frac{n}{2}0} \rangle} = \frac{1}{n^2} \sum_{m=0}^{n-1} (-1)^m \sum_{j=0}^{n-1} X_{mj} = \frac{\hat{X}_{\frac{n}{2}}}{n^2},$$

$$\frac{n^2}{2} \sum_{m=0}^{n-1} \langle X, E^{\frac{n}{2}\frac{n}{2}} \rangle = \frac{1}{n^2} \sum_{m=0}^{n-1} (-1)^m \sum_{j=0}^{n-1} X_{mj} = \frac{\hat{X}_{\frac{n}{2}}}{n^2}$$

e

$$\alpha_X^{\frac{n}{2}\frac{n}{2}} = \frac{\langle X, E^{\frac{n}{2}\frac{n}{2}} \rangle}{\langle E_{\frac{n}{2}\frac{n}{2}}, E_{\frac{n}{2}\frac{n}{2}} \rangle} = \frac{1}{n^2} \left[\sum_{m=0}^{n-1} \sum_{j=0}^{n-1} (-1)^{m+j} X_{mj} \right] = \frac{\hat{X}_{\frac{n}{2}\frac{n}{2}}}{n^2}.$$

Equivalentemente, no espaço de Fourier tem-se a projeção sobre o sub-espaço ortogonal aos vetores \hat{E}^{00} , $\hat{E}^{0\frac{n}{2}}$, $\hat{E}^{\frac{n}{2}0}$ e $\hat{E}^{\frac{n}{2}\frac{n}{2}}$,

$$\tilde{\hat{X}} = \hat{X} - \alpha_X^{00} \hat{E}^{00} - \alpha_X^{0\frac{n}{2}} \hat{E}^{0\frac{n}{2}} - \alpha_X^{\frac{n}{2}0} \hat{E}^{\frac{n}{2}0} - \alpha_X^{\frac{n}{2}\frac{n}{2}} \hat{E}^{\frac{n}{2}\frac{n}{2}}.$$

O ângulo entre duas projeções \tilde{X} e \tilde{Y} é dado por

$$\cos \sigma(\tilde{X}, \tilde{Y}) = \frac{\langle \tilde{X}, \tilde{Y} \rangle}{\|\tilde{X}\|_2 \|\tilde{Y}\|_2}$$

onde

$$\begin{split} \langle \tilde{X}, \tilde{Y} \rangle &= \langle X, Y \rangle - n^2 \alpha_X^{00} \alpha_Y^{00} - n^2 \alpha_X^{0\frac{n}{2}} \alpha_Y^{0\frac{n}{2}} - n^2 \alpha_X^{\frac{n}{2}0} \alpha_Y^{\frac{n}{2}0} - n^2 \alpha_X^{\frac{n}{2}\frac{n}{2}} \alpha_Y^{\frac{n}{2}\frac{n}{2}} \\ &= \frac{1}{n^4} [\langle \hat{X}, \hat{Y} \rangle - \hat{X}_{00} \hat{Y}_{00} - \hat{X}_{0\frac{n}{2}} \hat{Y}_{0\frac{n}{2}} - \hat{X}_{\frac{n}{2}0} \hat{Y}_{\frac{n}{2}0} - \hat{X}_{\frac{n}{2}\frac{n}{2}} \hat{Y}_{\frac{n}{2}\frac{n}{2}}] \end{split}$$

e

$$\begin{split} \|\tilde{X}\|_{2}^{2} &= \|X\|_{2}^{2} - n^{2}(\alpha_{X}^{00})^{2} - n^{2}(\alpha_{X}^{0\frac{n}{2}})^{2} - n^{2}(\alpha_{X}^{\frac{n}{2}0})^{2} - n^{2}(\alpha_{X}^{\frac{n}{2}\frac{n}{2}})^{2} \\ &= \frac{1}{n^{4}} [\|\hat{X}\|_{2}^{2} - (\hat{X}_{00})^{2} - (\hat{X}_{0\frac{n}{2}})^{2} - (\hat{X}_{\frac{n}{2}0})^{2} - (\hat{X}_{\frac{n}{2}\frac{n}{2}})^{2}] \end{split}$$

Sejam

$$g_i = \begin{bmatrix} X_{n-10} & X_{n-11} & \dots & X_{n-1n-1} \\ \vdots & & & \\ X_{10} & X_{11} & \dots & X_{1n-1} \\ X_{00} & X_{01} & \dots & X_{0n-1} \end{bmatrix},$$

$$\hat{g}_{i} = \begin{bmatrix} \mu_{10}\overline{\delta}_{10}, & \mu_{1n-1}\overline{\delta}_{1n-1} & \dots & \mu_{1\frac{n}{2}}\overline{\delta}_{1\frac{n}{2}} & \mu_{1\frac{n}{2}-1}\overline{\delta}_{1\frac{n}{2}-1} & \dots & \mu_{11}\overline{\delta}_{11} \\ \vdots & & & \\ \mu_{\frac{n}{2}-10}\overline{\delta}_{\frac{n}{2}-10} & \mu_{\frac{n}{2}-1n-1}\overline{\delta}_{\frac{n}{2}-1n-1} & \dots & \mu_{\frac{n}{2}-1\frac{n}{2}}\overline{\delta}_{\frac{n}{2}-1\frac{n}{2}} & \mu_{\frac{n}{2}-1\frac{n}{2}-1}\overline{\delta}_{\frac{n}{2}-1\frac{n}{2}-1} & \dots & \mu_{\frac{n}{2}-11}\overline{\delta}_{\frac{n}{2}-11} \\ \mu_{\frac{n}{2}0}\overline{\epsilon}_{\frac{n}{2}0} & \mu_{\frac{n}{2}1}\delta_{\frac{n}{2}1} & \dots & \mu_{\frac{n}{2}\frac{n}{2}}\overline{\epsilon}_{\frac{n}{2}\frac{n}{2}} & \mu_{\frac{n}{2}\frac{n}{2}-1}\overline{\delta}_{\frac{n}{2}\frac{n}{2}-1} & \dots & \mu_{\frac{n}{2}\frac{n}{2}1}\overline{\delta}_{\frac{n}{2}1} \\ \vdots & & & & \\ \mu_{10}\delta_{10} & \mu_{11}\delta_{11} & \dots & \mu_{1\frac{n}{2}}\delta_{1\frac{n}{2}} & \mu_{1\frac{n}{2}+1}\delta_{1\frac{n}{2}+1} & \dots & \mu_{1n-1}\delta_{1n-1} \\ \mu_{00}\overline{\epsilon}_{00} & \mu_{01}\varphi_{01} & \dots & \mu_{0\frac{n}{2}}\overline{\epsilon}_{0\frac{n}{2}} & \mu_{0\frac{n}{2}-1}\overline{\delta}_{0\frac{n}{2}-1} & \dots & \mu_{01}\overline{\delta}_{01} \end{bmatrix}$$
$$g'_{i-1} = \begin{bmatrix} Y_{n-10} & Y_{n-11} & \dots & Y_{n-1n-1} \\ \vdots & & & \\ Y_{10} & Y_{11} & \dots & Y_{1n-1} \\ Y_{00} & Y_{01} & \dots & Y_{0n-1} \end{bmatrix}$$

e

$$\hat{g}_{i-1}' = \begin{bmatrix} \lambda_{10}\overline{\varphi}_{10}, & \lambda_{1n-1}\overline{\varphi}_{1n-1} & \dots & \lambda_{1\frac{n}{2}}\overline{\varphi}_{1\frac{n}{2}} & \lambda_{1\frac{n}{2}-1}\overline{\varphi}_{1\frac{n}{2}-1} & \dots & \lambda_{11}\overline{\varphi}_{11} \\ \vdots & & & & \\ \lambda_{\frac{n}{2}-10}\overline{\varphi}_{\frac{n}{2}-10} & \lambda_{\frac{n}{2}-1n-1}\overline{\varphi}_{\frac{n}{2}-1n-1} & \dots & \lambda_{\frac{n}{2}-1\frac{n}{2}}\overline{\varphi}_{\frac{n}{2}-1\frac{n}{2}} & \lambda_{\frac{n}{2}-1\frac{n}{2}-1}\overline{\varphi}_{\frac{n}{2}-1\frac{n}{2}-1} & \dots & \lambda_{\frac{n}{2}-11}\overline{\varphi}_{\frac{n}{2}-1} \\ \lambda_{\frac{n}{2}0}\varepsilon_{\frac{n}{2}0} & \lambda_{\frac{n}{2}1}\varphi_{\frac{n}{2}1} & \dots & \lambda_{\frac{n}{2}\frac{n}{2}}\varepsilon_{\frac{n}{2}\frac{n}{2}} & \lambda_{\frac{n}{2}\frac{n}{2}-1}\overline{\varphi}_{\frac{n}{2}\frac{n}{2}-1} & \dots & \lambda_{\frac{n}{2}1}\overline{\varphi}_{\frac{n}{2}1} \\ \vdots & & & & \\ \lambda_{10}\varphi_{10} & \lambda_{11}\varphi_{11} & \dots & \lambda_{1\frac{n}{2}}\varphi_{1\frac{n}{2}} & \lambda_{1\frac{n}{2}+1}\varphi_{1\frac{n}{2}+1} & \dots & \lambda_{1n-1}\varphi_{1n-1} \\ \lambda_{00}\varepsilon_{00} & \lambda_{01}\varphi_{01} & \dots & \lambda_{0\frac{n}{2}}\varepsilon_{0\frac{n}{2}} & \lambda_{0\frac{n}{2}-1}\overline{\varphi}_{0\frac{n}{2}-1} & \dots & \lambda_{01}\overline{\varphi}_{01} \\ \end{bmatrix}$$

onde μ_{mj} , δ_{mj} e γ_{mj} representam, respectivamente, amplitudes na *i*-ésima iteração, fases na *i*-ésima iteração e fases na iteração i - 1. Os λ_{mj} são os dados.

Tem-se
$$\alpha_{g_{i}}^{00} = \frac{\mu_{00}}{n^{2}}, \quad \alpha_{g_{i}}^{0\frac{n}{2}} = \frac{\mu_{0\frac{n}{2}}\overline{\varepsilon}_{0\frac{n}{2}}}{n^{2}}, \quad \alpha_{g_{i}}^{\frac{n}{2}} = \frac{\mu_{\frac{n}{2}}\overline{\varepsilon}_{\frac{n}{2}}}{n^{2}}, \quad \alpha_{g_{i}}^{\frac{n}{2}\frac{n}{2}} = \frac{\mu_{\frac{n}{2}}\overline{\varepsilon}_{\frac{n}{2}}}{n^{2}}, \quad \alpha_{g_{i-1}}^{00} = \frac{\lambda_{00}}{n^{2}}, \quad \alpha_{g_{i-1}}^{\frac{n}{2}} = \frac{\lambda_{00}}{n^{2}}, \quad \alpha_{g_{i-1}}^{\frac{n}{2}\frac{n}{2}} = \frac{\lambda_{00}}{n^{2}}, \quad \alpha_{g_{i-1}}^{\frac{n}{2}\frac{n}{2}\frac{n}{2}}, \quad \alpha_{g_{i-1}}^{00} = \frac{\lambda_{00}}{n^{2}}, \quad \alpha_{g_{i-1}}^{\frac{n}{2}\frac{n}{2}\frac{n}{2}}, \quad \alpha_{g_{i-1}}^{\frac{n}{2}\frac{n}{2}\frac{n}{2}}, \quad \alpha_{g_{i-1}}^{\frac{n}{2}\frac{n}{2}\frac{n}{2}\frac{n}{2}}, \quad \alpha_{g_{i-1}}^{\frac{n}{2}\frac{n}{2}\frac{n}{2}\frac{n}{2}}, \quad \alpha_{g_{i-1}}^{\frac{n}{2}\frac{n}{2}\frac{n}{2}\frac{n}{2}}, \quad \alpha_{g_{i-1}}^{\frac{n}{2}\frac{n}{2}\frac{n}{2}\frac{n}{2}\frac{n}{2}}, \quad \alpha_{g_{i-1}}^{\frac{n}{2}\frac{n}{2}\frac{n}{2}\frac{n}{2}\frac{n}{2}}, \quad \alpha_{g_{i-1}}^{\frac{n}{2}\frac{n}{2}\frac{n}{2}\frac{n}{2}\frac{n}{2}}, \quad \alpha_{g_{i-1}}^{\frac{n}{2}\frac{n}{2}\frac{n}{2}\frac{n}{2}\frac{n}{2}}, \quad \alpha_{g_{i-1}}^{\frac{n}{2}\frac{n}{2}\frac{n}{2}\frac{n}{2}\frac{n}{2}\frac{n}{2}}, \quad \alpha_{g_{i-1}}^{\frac{n}{2}\frac{n}{2}\frac{n}{2}\frac{n}{2}\frac{n}{2}\frac{n}{2}\frac{n}{2}\frac{n}{2}\frac{n}{2}}, \quad \alpha_{g_{i-1}}^{\frac{n}{2}\frac{n}{$$

Lema 7 Se a seqüência $\{g_i\}$ não é limitada e $g_0 \ge 0$, então existem uma subseqüência $\{g_{i_p}\}$ com $i_p \in \mathbb{N}$, $p = 1, 2, ..., i_p < i_{p+1}$, um subconjunto $\Omega \subset \{0, 2, ..., n-1\} \times \{0, 2, ..., n-1\}$ e uma constante K tais que:

1) Para $(m, j) \in \Omega$,

$$\begin{cases} g_{i_p,mj} \to +\infty, & g'_{i_p-1,mj} < 0 \\ g_{i_p-1,mj} = g_{i_p,mj} + \beta g'_{i_p-1,mj} \end{cases}$$

2) Para $(m, j) \notin \Omega$

$$\begin{cases} g_{i_p,mj} < K \\ g_{i_p-1,mj} < K \end{cases}$$
3) Para todo (m, j) :

$$|g'_{i_p-1,mj}| < K$$

4) Os sinais de

$$\mu_{0\frac{n}{2}}\overline{\varepsilon}_{0\frac{n}{2}} = \frac{\sum_{m=0}^{n-1}\sum_{j=0}^{n-1}(-1)^{j}X_{mj}}{n^{2}}, \quad \mu_{\frac{n}{2}0}\overline{\varepsilon}_{\frac{n}{2}0} = \frac{\sum_{m=0}^{n-1}(-1)^{m}\sum_{j=0}^{n-1}X_{mj}}{n^{2}}$$
$$\mu_{\frac{n}{2}\frac{n}{2}}\overline{\varepsilon}_{\frac{n}{2}\frac{n}{2}} = \frac{\sum_{m=0}^{n-1}\sum_{j=0}^{n-1}(-1)^{m+j}X_{mj}}{n^{2}}$$

são fixos;

е

Prova:

Como a seqüência $\{g_i\}$ não é limitada, $\{||g_i||_1\}$ não é limitada e pode-se escolher uma subseqüência tal que $||g_{i-1}||_1 < ||g_i||_1 \rightarrow +\infty$. Considerando (m,j)=(0,0), se existe K_{00} tal que $|g_{i,00}| < K_{00}$ para todo i declara-se que $(0,0) \notin \Omega$; se não, pode-se escolher uma subseqüência da subseqüência anterior para a qual $|g_{i,00}| \rightarrow +\infty$ e declarar $(0,0) \in \Omega$. Fazendo isso, indutivamente, para todo $(m,j) \in \{0,1,2,...n-1\} \times \{0,1,2,...n-1\}$ e escolhendo K suficientemente grande o resultado é uma subseqüência que satisfaz (1) e (2). A é compacto logo a condição (3) é satisfeita. Como a subseqüência obtida pertence a no máximo oito toróides é possivel escolher uma subseqüência da mesma pertencendo a um único toróide e portanto satisfazendo 4). No que segue, sempre que citarmos $g_{i,mj}$ estaremos nos referindo à subseqüência do lema anterior.

Lema 8 Para

$$\overline{\lambda} = \max_{(m,j)\notin L} \{\lambda_{mj}\} \quad e \quad \underline{\lambda} = \min_{(m,j)\notin L} \{\lambda_{mj}\}$$

com L={ $(0,0), (0, \frac{n}{2}), (\frac{n}{2}, 0), (\frac{n}{2}, \frac{n}{2})$ } tem-se:

$$\cos\sigma(\tilde{g}_{i-1}, \tilde{g}'_{i-1}) \ge \frac{\underline{\lambda}}{\sqrt{[(n-1)^2 - 4]}\overline{\lambda}}$$

Prova:

Verifica-se que

$$\|\tilde{g}_{i-1}'\|_{2}^{2} = \frac{1}{n^{4}} \sum_{(m,j)\notin L} \lambda_{mj}^{2}, \quad \|\tilde{g}_{i-1}\|_{2}^{2} = \frac{1}{n^{4}} \sum_{(m,j)\notin L} \mu_{mj}^{2}, \quad e \quad \langle \tilde{g}_{i-1}', \tilde{g}_{i-1} \rangle = \frac{1}{n^{4}} \sum_{(m,j)\notin L} \lambda_{mj} \mu_{mj}.$$

Por estimativas simples,

$$\cos \sigma(\tilde{g}_{i-1}, \tilde{g}'_{i-1}) = \frac{\langle \tilde{g}'_{i-1}, \tilde{g}_{i-1} \rangle}{\|\tilde{g}'_{i-1}\|_2 \|\tilde{g}_{i-1}\|_2}$$
$$\geq \frac{\lambda}{\sqrt{[(n-1)^2 - 4]\lambda}} \cdot \frac{\sum_{\substack{(m,j) \notin L}} \mu_{mj}}{\left[\sum_{(m,j) \notin L} \mu_{mj}^2\right]^{1/2}} \geq \frac{\lambda}{\sqrt{[(n-1)^2 - 4]\lambda}}.$$

C.Q.D.

Lema 9 Para i muito grande,

$$\cos \sigma(\tilde{g}'_{i-1}, \tilde{g}_i) < -\frac{\lambda_{00}}{\left[\sum_{(m,j)\notin L} \lambda_{mj}^2\right]^{1/2}}.$$

Prova:

Temos que

$$\langle \tilde{g}'_{i-1}, \tilde{g}_i \rangle = \langle g'_{i-1}, g_i \rangle - \frac{1}{n^2} \left[\lambda_{00} \mu_{00} - \lambda_{0\frac{n}{2}} \varepsilon_{0\frac{n}{2}} \mu_{0\frac{n}{2}} \overline{\varepsilon}_{0\frac{n}{2}} - \lambda_{\frac{n}{2}0} \varepsilon_{\frac{n}{2}0} \mu_{\frac{n}{2}0} \overline{\varepsilon}_{\frac{n}{2}0} - \lambda_{\frac{n}{2}\frac{n}{2}} \varepsilon_{\frac{n}{2}\frac{n}{2}} \mu_{\frac{n}{2}\frac{n}{2}} \overline{\varepsilon}_{\frac{n}{2}\frac{n}{2}} \right]$$

e

$$\|\tilde{g}_i\|_2^2 = \sum_{m=0}^{n-1} \sum_{j=0}^{n-1} X_{mj}^2 - \frac{1}{n^2} \left(\sum_{m=0}^{n-1} \sum_{j=0}^{n-1} X_{mj} \right)^2 - \frac{1}{n^2} \left[\left(\sum_{m=0}^{n-1} \sum_{j=0}^{n-1} (-1)^j X_{mj} \right)^2 + \left(\sum_{m=0}^{n-1} (-1)^m \sum_{j=0}^{n-1} X_{mj} \right)^2 + \left(\sum_{m=0}^{n-1} \sum_{j=0}^{n-1} (-1)^m X_{mj} \right)^2 \right].$$

Logo

$$\begin{split} \cos \sigma(\tilde{g}_{i-1}', \tilde{g}_i) &= \frac{\langle \tilde{g}_{i-1}', \tilde{g}_i \rangle}{\|\tilde{g}_{i-1}'\|_2 \|\tilde{g}_i\|_2} \leq \frac{\langle g_{i-1}', g_i \rangle - \frac{\lambda_{00}}{n^2} \mu_{00}}{\|\tilde{g}_{i-1}'\|_2 \|\tilde{g}_i\|_2} \\ &= \frac{\langle g_{i-1}', g_i \rangle}{\|\tilde{g}_{i-1}'\|_2 \|\tilde{g}_i\|_2} - \frac{\frac{\lambda_{00}}{n^2} \mu_{00}}{\|\tilde{g}_{i-1}'\|_2 \|\tilde{g}_i\|_2} \\ &\leq \frac{\langle g_{i-1}', g_i \rangle}{\|\tilde{g}_{i-1}'\|_2 \|\tilde{g}_i\|_2} - \frac{\frac{\lambda_{00}}{n^2} \sum_{m=0}^{n-1} \sum_{j=0}^{n-1} X_{mj}}{\|\tilde{g}_{i-1}'\|_2 \left[\sum_{m=0}^{n-1} \sum_{j=0}^{n-1} X_{mj}^2\right]^{1/2}} \\ &\leq \frac{\sum_{m=0}^{n-1} Y_{mj} X_{mj} + \sum_{(m,j) \notin \Omega} Y_{mj} X_{mj}}{\|\tilde{g}_{i-1}'\|_2 \|\tilde{g}_i\|_2} - \frac{\frac{\lambda_{00}}{n^2}}{\frac{1}{n^2} \left[\sum_{(m,j) \notin L} \lambda_{mj}^2\right]^{1/2}}. \end{split}$$

Como,

$$\sum_{(m,j)\in\Omega} Y_{mj} X_{mj} \to -\infty, \quad e \quad \sum_{(m,j)\notin\Omega} Y_{mj} X_{mj}$$

é limitada, para i suficientemente grande,

Lema 10

$$\lim_{i \to \infty} \frac{\|\tilde{g}_{i-1}\|_2}{\|\tilde{g}_i\|_2} = 1.$$

Prova:

Pela construção da subseqüência no Lema 7 tem-se 0 < $\omega = |\Omega| < n^2$. Podemos escrever:

$$\|\tilde{g}_{i}\|_{2}^{2} = \sum_{(m,j)} X_{mj}^{2} - \frac{1}{n^{2}} \left(\sum_{(m,j)} X_{mj} \right)^{2} - \frac{1}{n^{2}} \left[\left(\sum_{m=0}^{n-1} \sum_{j=0}^{n-1} (-1)^{j} X_{mj} \right)^{2} + \left(\sum_{m=0}^{n-1} (-1)^{m} \sum_{j=0}^{n-1} X_{mj} \right)^{2} + \left(\sum_{m=0}^{n-1} \sum_{j=0}^{n-1} (-1)^{m+j} X_{mj} \right)^{2} \right] = A + B$$

onde

e

$$A = \sum_{m \in \Omega} X_{mj}^2 - \frac{1}{n^2} \left(\sum_{m \in \Omega} X_m \right)^2$$
$$B = \sum_{(m,j) \notin \Omega} X_{mj}^2 - \frac{1}{n^2} \sum_{(m,j) \notin \Omega} X_{mj} \left(\sum_{(m,j) \notin \Omega} X_{mj} + 2 \sum_{(m,j) \in \Omega} X_{mj} \right)$$
$$- \frac{1}{n} \left[\left(\sum_{m=0}^{n-1} \sum_{j=0}^{n-1} (-1)^j X_{mj} \right)^2 + \left(\sum_{m=0}^{n-1} (-1)^m \sum_{j=0}^{n-1} X_{mj} \right)^2 + \left(\sum_{m=0}^{n-1} \sum_{j=0}^{n-1} (-1)^{m+j} X_{mj} \right)^2 \right].$$

Vamos provar que para i muito grande,

$$B \ll \frac{\sum_{(m,j)\in\Omega} X_{mj}^2}{n} < A.$$

Em primeiro lugar,

$$A = \frac{1}{n^2} \left\{ \omega \sum_{(m,j)\in\Omega} X_{mj}^2 - \left(\sum_{(m,j)\in\Omega} X_{mj} \right)^2 \right\} + \frac{n^2 - \omega}{n^2} \sum_{(m,j)\in\Omega} X_{mj}^2$$
$$\geq \frac{\sum_{(m,j)\in\Omega} X_{mj}^2}{n^2}$$

pois,

$$\omega \sum_{(m,j)\in\Omega} X_{mj}^2 - (\sum_{(m,j)\in\Omega} X_{mj})^2 \ge 0.$$

Pela limitação dos elementos e por estimativas simples, baseadas no Lema 7,

$$B \le (n^2 - \omega)K^2 + \frac{(n^2 - \omega)}{n^2}K\left((n^2 - \omega)K + 2\sum_{(m,j)\in\Omega} X_{mj}\right) = L_0 + L_1\sum_{(m,j)\in\Omega} X_{mj}$$

Como, para i muito grande,

$$\sum_{(m,j)\in\Omega} X_{mj} \ll \frac{1}{n^2} \sum_{(m,j)\in\Omega} X_{mj}^2$$

vem que

$$B \ll \frac{1}{n^2} \sum_{m \in \Omega} X_m^2 \le A.$$

Definindo $Z_m = g_{i-1,m} = X_m + \beta Y_m$ pode-se escrever

$$\|\tilde{g}_{i-1}\|_{2}^{2} = \sum_{m=0}^{n-1} \sum_{j=0}^{n-1} Z_{mj}^{2} - \frac{1}{n^{2}} \left(\sum_{m=0}^{n-1} \sum_{j=0}^{n-1} Z_{mj} \right)^{2}$$
$$-\frac{1}{n^{2}} \left[\left(\sum_{m=0}^{n-1} \sum_{j=0}^{n-1} (-1)^{j} Z_{mj} \right)^{2} + \left(\sum_{m=0}^{n-1} (-1)^{m} \sum_{j=0}^{n-1} Z_{mj} \right)^{2} + \left(\sum_{m=0}^{n-1} \sum_{j=0}^{n-1} (-1)^{m+j} Z_{mj} \right)^{2} \right] = A + C$$

onde

$$C = \beta \sum_{(m,j)\in\Omega} Y_{mj} \left(\beta Y_{mj} + 2X_{mj}\right) + \sum_{(m,j)\notin\Omega} Z_{mj}^2$$
$$-\frac{\beta}{n^2} \left(\sum_{(m,j)\in\Omega} Y_{mj}\right) \left(\beta \sum_{(m,j)\in\Omega} Y_{mj} + 2\sum_{(m,j)\in\Omega} X_{mj}\right) - \frac{1}{n^2} \left(\sum_{(m,j)\in\Omega} Z_{mj}\right)^2$$

$$-\frac{1}{n^2} \left[\left(\sum_{m=0}^{n-1} \sum_{j=0}^{n-1} (-1)^j Z_{mj} \right)^2 + \left(\sum_{m=0}^{n-1} (-1)^m \sum_{j=0}^{n-1} Z_{mj} \right)^2 + \left(\sum_{m=0}^{n-1} \sum_{j=0}^{n-1} (-1)^{m+j} Z_{mj} \right)^2 \right].$$

Por estimativas similares as feitas no denominador,

$$C \leq M_0 + M_1 \sum_{(m,j)\in\Omega} X_{mj} \ll \frac{1}{n^2} \sum_{(m,j)\in\Omega} X_{mj}^2.$$

Consequentemente,

$$\lim_{i \to \infty} \frac{\|\tilde{g}_{i-1}\|_2}{\|\tilde{g}_i\|_2} = \lim_{i \to \infty} \sqrt{\frac{A+C}{A+B}} = 1.$$

C.Q.D.

Lema 11

$$\lim_{i \to \infty} \cos \sigma(\tilde{g}_{i-1}, \tilde{g}_i) = 1.$$

Prova:

$$\lim_{i \to \infty} \frac{\langle \tilde{g}_{i-1}, \tilde{g}_i \rangle}{\|\tilde{g}_{i-1}\|_2 \|\tilde{g}_i\|_2}$$

$$\geq \lim_{i \to \infty} \frac{\|\tilde{g}_i\|_2^2 - 5K\beta \sum_{m=0}^{n-1} \sum_{j=0}^{n-1} X_{mj}}{\|\tilde{g}_i\|_2^2} \cdot \lim_{i \to \infty} \frac{\|\tilde{g}_i\|_2}{\|\tilde{g}_{i-1}\|_2}$$

$$= \lim_{i \to \infty} \frac{\|\tilde{g}_i\|_2^2 - 5K\beta \sum_{m=0}^{n-1} \sum_{j=0}^{n-1} X_{mj}}{\|\tilde{g}_i\|_2^2}.$$

Logo,

$$\lim_{i \to \infty} \cos \sigma(\tilde{g}_{i-1}, \tilde{g}_i) = 1$$

Com estes lemas a demonstração do Teorema 1 é exatamente a mesma que para recuperação de sinais unidimensionais.

5.1.4 Input-Output

A prova de que este algoritmo é limitado é análoga à feita para o algoritmo híbrido. Basta observar que aqui o Lema 1 também é verdadeiro e que no Lema 2 se a seqüência obtida pelo algoritmo não é limitada existe uma subseqüência não limitada igual a encontrada para o algoritmo híbrido. Conseqüentemente todas a outras provas feitas são iguais.

5.2 Pontos Fixos

Tal como descrito em [33], existem dois tipos de ambiguidades do problema da recuperação da fase: aquelas que dependem do problema, que chamaremos de intrínsecas, e aquelas que são causadas pelo algoritmo usado para resolvê-lo, as não intrínsecas. Evidentemente, as ambiguidades não intrínsecas são as que estão relacionadas com os pontos fixos dos algoritmos que não são solução, ou com os possíveis ciclos. Nesta seção analisamos e comentamos brevemente os pontos fixos de cada um dos algoritmos. Aqui estaremos considerando que no espaço tempo as restrições são a não negatividade e o suporte da imagem.

5.2.1 Error-Reduction

Para este algoritmo,

$$g_{k}(x) = g_{k+1}(x) = \begin{cases} g'_{k}(x) & \text{se } x \in \gamma \\ 0 & \text{se } x \notin \gamma \end{cases}$$

se, e somente se, $g'_k(x) = g_k(x)$ quando $x \in \gamma$ e $g'_k(x)$ é qualquer em outro caso. Como uma solução do problema de fase satisfaz $g_k(x) = g'_k(x)$ para todo x, conclui-se que (se $g'_k(x) \neq 0$ para $x \notin \gamma$) o algoritmo pode estagnar gerando duas seqüências associadas a pontos fixos diferentes, um satisfazendo a não negatividade e outro satifazendo os dados.

5.2.2 Output-Output

Aqui, para $\beta \neq 0$ e $\beta \neq 1$ (que leva ao ER),

$$g_{k}(x) = g_{k+1}(x) = \begin{cases} g'_{k}(x) & \text{se } x \in \gamma \\ g'_{k}(x) - \beta g'_{k}(x) & \text{se } x \notin \gamma \end{cases}$$

se, e somente se, $g'_k(x) = g_k(x)$ quando $x \in \gamma$ e $g'_k(x) = \frac{1}{1-\beta}g_k(x)$ em outro caso. Novamente, quando não se trata de uma solução, temos duas seqüências associadas a pontos fixos diferentes, como no caso anterior. Tal como já observado, o algoritmo OO é apenas uma relaxação do ER quando se projeta sobre as restrições, e portanto espera-se um comportamento similar do ponto de vista das estagnações.

5.2.3 Input-Output

Como o passo deste algoritmo é dado por

$$g_{k+1}(x) = \begin{cases} g_k(x) & \text{se } x \in \gamma \\ g_k(x) - \beta g'_k(x) & \text{se } x \notin \gamma \end{cases},$$

para $\beta \neq 0$, g_k é ponto fixo se, e somente se, $g'_k(x) = 0$ quando $x \notin \gamma$. Assim $g'_k(x)$ é uma solução.

5.2.4 Input-Output Híbrido

Neste algoritmo, como $\beta \neq 0$, g_k é ponto fixo do algoritmo se, e somente se, $g'_k(x) = g_k(x)$ quando $x \in \gamma$ e $g'_k(x) = 0$ em outro caso, mas então nos encontramos na situação anterior, $g'_k(x)$ é solução, e portanto o algoritmo não tem pontos fixos independentes de uma solução.

Capítulo 6

UM MÉTODO AUTOMÁTICO PARA SUPERAR A ESTAGNAÇÃO

O principal problema dos algoritmos para recuperar a fase da TF dada a amplitude é que muitas vezes eles estagnam em pontos que não são soluções. Como já comentamos, este comportamento tem sido atribuido essencialmente a dois motivos: (1) A existência de pontos bastante diferentes da solução mas que quase satisfazem as restrições, as chamadas ambiguidades intrínsecas; e (2) a existência de pontos fixos, ou pontos de estagnação, dos algoritmos, as ambiguidades não intrínsecas. Para resolver o problema gerado por (2), várias técnicas tem sido propostas. Em [19], as técnicas são dependentes do tipo de estagnação, que, no caso, é apenas detectada visualmente, tal como descrito no Capítulo 2. Outra forma, proposta por Fienup [16], para sair da estagnação é mudar de algoritmo quando isto é detectado, porém esta técnica nem sempre funciona. Em [39] e [9], outra metodologia para resolver o problema é sugerida, mas supõe algumas hipóteses sobre a solução que não são necessariamente verdadeiras. Em [33], é sugerida uma combinação de vários métodos de otimização usando uma função objetivo muito particular.

A nossa proposta, em princípio, é baseada na análise da geometria do problema descrita no Capítulo 4 e no possível comportamento do algoritmo ER quando fica estagnado.

Em [29] é apresentada uma figura descrevendo a geometria possível de um ponto de estagnação para ER (ou OO). A idéia é reproduzida na Figura 6.1.



Figura 6.1: Geometria dos Pontos de Estagnação I

Tomando a figura acima como ponto de partida, pareceria que uma direção correta para sair do ponto de estagnação seria 'perpendicular' à direção determinada pelos dois pontos da iteração pertencentes, respectivamente, aos dois conjuntos determinados pelas restrições. Sem entrar no mérito da noção de perpendicularidade, no caso do nosso interesse, o aspecto geométrico é completamente diferente. Uma completa representação da caracterização geométrica não é possível por causa da dimensão do espaço que contém os toróides, mas podemos concluir muitas coisas a partir dos resultados do Capítulo 4. A primeira observação, obtida com a ajuda da experiência numérica, é que os pontos de oscilação do ER ocorrem entre um toróide e o ortante não negativo. Se pensamos o toróide como uma superfície rodeando um conjunto convexo fechado, e se o algoritmo estagna entre o toróide e uma aresta do ortante não negativo, é razoável recomeçar usando um ponto localizado 'do outro lado' do toróide. Uma figura que melhor aproxima a situação descrita é a seguinte:



Figura 6.2: Geometria dos Pontos de Estagnação II

Baseados na dedução anterior apresentamos o novo método para superar a estagnação na próxima seção.

6.1 O Novo Método

O método consiste no que segue:

Suponhamos que o algoritmo estagnou numa imagem g_k cuja projeção sobre o conjunto de

toróides é g'_k . Obtém-se uma nova estimativa inicial para o algoritmo, refletindo g'_k através do centro do toróide ao qual esta pertence, isto é, $g^n_o = -g'_k + 2c \mod c$ representando o centro do toróide. Este novo ponto inicial pertence ao mesmo toróide ao qual pertence g'_k , portanto satisfaz as restrições do domínio de Fourier e, a menos que seja solução, não satisfaz as restrições do domínio da imagem. O cálculo deste novo ponto inicial é muito simples e consiste apenas numa mudança de sinal nos elementos da matriz da TF que sejam diferentes daqueles (reais) que definem o toróide.

No próximo capítulo apresentamos e comentamos as experiências numéricas realizadas com os novos métodos.

Capítulo 7

EXPERIMENTOS

O objetivo deste capítulo é descrever os experimentos realizados no desenvolvimento do trabalho e mostrar alguns dos resultados.

7.1 Linguagem e pacotes utilizados

Todos os programas que escrevemos utilizam a linguagem Fortran e podem ser executados na SUN. Para visualizar e imprimir as imagens utilizamos o display do **Snark93** e o **xv**. O **xv** está disponivel na SUN e o **Snark93** é um programa para reconstrução de imagens através de projeções, desenvolvido pelo grupo de processamento de imagens médicas do Departamento de Radiologia na Universidade da Pennsylvania.

7.2 Imagens e seus formatos

Utilizamos imagens com (ou contidas num) suporte quadrado de lado contendo um número de pixels igual a uma potência de dois para que a transformada rápida de Fourier pudesse ser utilizada. Estas imagens estão armazenadas na forma matricial.

7.3 Implementação dos algoritmos

Os algoritmos de Fienup (error-reduction, output-output, input-output e híbrido) assim como alguns dos outros propostos foram programados para recuperar sinais e recuperar imagens. Fizemos muitos testes com sinais unidimensionais mas, nos fixamos na recuperação de imagens.

Como em cada iteração de todos os algoritmos é preciso calcular a TF e sua inversa utilizamos uma subrotina que calcula a transformada rápida de Fourier.

Observamos que o algoritmo IOH foi programado de duas formas:

1) em sua versão original (que chamaremos de Versão 1), onde as restrições de suporte fazem parte da definição do passo, e

2) incluindo as restrições de suporte na forma de uma projeção após o passo (Versão 2).

No caso de recuperação de imagens, a janela utilizada foi um quadrado, cujo comprimento do lado é duas o comprimento do lado do quadrado que contém a imagem. Tal janela contém o suporte da autocorrelação (único dado sobre o suporte sempre disponível).

7.3.1 Dados

Utilizamos como informações sobre a imagem a ser recuperada a amplitude da transformada de Fourier, não negatividade da função que representa a imagem e restrições de suporte no domínio da imagem. Como suporte utilizamos, por exemplo, o suporte da autocorrelação ou, dividindo a janela em quatro quadrados iguais, o quadrado do canto inferior esquerdo [11].

Fizemos testes considerando um ruído Gaussiano de 10% nos dados. Os resultados obtidos foram semelhantes aos do caso sem ruído. Assim apresentamos somente exemplos

cujos dados não contém ruído.

7.3.2 Imagens iniciais

Como estimativas iniciais, em geral, utilizamos uma função gerada por números aleatórios com suporte igual ao da autocorrelação, o centro do toróide em que se encontra a imagem, ou alguma variação destas possibilidades.

7.3.3 Testes para parada e acompanhamento dos programas

Considerando-se que f representa a imagem a ser recuperada e g_k a estimativa da imagem na iteração k, com F e G_k suas respectivas transformadas de Fourier, em cada iteração foram calculadas as seguintes medidas:

$$ET = \left[\frac{\sum_{x}(g_{k}(x) - f(x))^{2}}{\sum_{u} |F(u)|^{2}}\right]^{1/2},$$
$$E0 = \left[\frac{\sum_{x \notin \gamma}(g'_{k}(x))^{2}}{\sum_{u} |F(u)|^{2}}\right]^{1/2}$$

e

$$EF = \left[\frac{\sum_{u} (|G_{k}(u)| - |F(u)|)^{2}}{\sum_{u} |F(u)|^{2}}\right]^{1/2}$$

ET representa a distância à imagem a ser recuperada (estimativa que não pode ser calculada na prática e que em geral não faz sentido pois os algoritmos podem convergir para associadas triviais da imagem), E0 representa a distância de g'_k ao conjunto das imagens que satisfazem as restrições de não-negatividade e de suporte e EF a distância de g_k ao conjunto das imagens cuja transformada de Fourier tem a amplitude dada.

Incluimos no algoritmo um teste de parada envolvendo EF ($EF < 1 \times 10^{-5})$ e outro

baseado no número de iterações. Em geral os programas foram interrompidos pelo segundo teste (este é um dos problemas a ser resolvido).

7.4 Algoritmos de Fienup utilizando informação sobre a caracterização

Testamos os algoritmos de Fienup restringindo as projeções do espaço de Fourier a um dos oito toróides que definem as imagens não negativas com amplitude da transformada de Fourier dada. Nos exemplos que consideramos esta restrição não levou a uma sensível melhora dos algoritmos. Observamos também, que nesses exemplos, as estimativas obtidas nos algoritmos sem esta restrição em geral permanecem num mesmo toróide.

7.5 Resultados

As Figuras 1 e 2 contém as imagens que estaremos tentando recuperar, utilizando a amplitude de sua TF. As três próximas Figuras contém as imagens que foram utilizadas como estimativas iniciais para os algoritmos, nos exemplos apresentados aqui.

No Exemplo 1, o objetivo é recuperar a imagem 1 utilizando o algoritmo das projeções generalizadas modificado (proposto no Capítulo 3). Para o resultado da Figura 7.6 foram utilizados os parâmetros $\lambda_1 = 0.99$ e $\lambda_2 = 2$ nas três primeiras iterações do algoritmo e em seguida iguais, respectivamente, aos valores a_1 e a_2 (limitantes superiores para os parâmetros que aparecem na demonstração de Levi e Stark, Seção 3.2). A Figura 7.7 mostra o resultado da Versão 2 do algoritmo IOH aplicado a mesma Imagem inicial que para o PGM. O parâmetro utilizado foi $\beta = 1$ (que para esta versão tem sido uma boa opção). A Versão 1 do algoritmo IOH, com este mesmo parâmetro, passa pelas estagnações (inclusive pela solução) sem se fixar.

Os Exemplos 2, 3 e 4 mostram a eficiência do método proposto no Capítulo 6 no caso de estagnações do algoritmo ER em imagens com suporte truncado, do tipo gêmeas ou até estagnações consideradas intrínsicas. Nos dois primeiros exemplos o objetivo é recuperar a Imagem 1 e no último a Imagem 2. A aplicação das reflexões fez com que o algoritmo convergisse para a imagem (ou uma associada). As medidas ET, E0 e EF são as descritas na Subseção 7.3.3.

A Figura 7.8 mostra um exemplo de imagem que classificamos como gêmea. No Exemplo 4, Figura 7.18, pode-se ver a ocorrência de uma estagnação do tipo suporte truncado.

Aplicamos a Versão 2 do algoritmo híbrido (caso para o qual provamos a limitação) para os mesmos exemplos da seção anterior. Ocorrem estagnações do mesmo tipo que para o error-reduction, embora com EF e EO menores. No Exemplo 4, considerandindo $\beta = 1$, o algoritmo híbrido teve aproximadamente o mesmo comportamento que o error-reduction, isto é após 100 iterações convergiu para imagem com suporte truncado da Figura 7.18. Depois de uma reflexão pelo centro do toróide e mais 100 iterações do algoritmo híbrido convergiu para uma imagem sem suporte truncado. As medidas EF e EO foram menores que para o error-reduction nas duas imagens obtidas e acontece uma pequena melhora na qualidade das imagens.

O Exemplo 5 descreve o comportamento desta versão do algoritmo IOH, utilizando $\beta = 1$, na recuperação da Imagem 1 com Imagem inicial 2.

Em todos os exemplos citados o suporte utilizado foi igual ao quadrado inferior esquerdo, equivalente a um quarto da janela.

Sempre que encontramos estagnações, testamos nosso método e recuperamos a imagem com no máximo três reflexões pelo centro.

Observamos que uma quantidade maior de iterações, que as consideradas nos exemplos,

não melhora significativamente a qualidade das imagems e que uma nova reflexão pode tirar da solução considerada razoável para uma outra estagnação. Novas reflexões levam novamente a uma solução.

7.5.1 Exemplos

Figura 7.1: Imagem 1



Figura 7.2: Imagem 2



Figura 7.3: Imagem Inicial 1



Figura 7.4: Imagem Inicial 2



Figura 7.5: Imagem inicial 3

Exemplo 1



Figura 7.6: Após 100 iterações do algoritmo PGM, considerando a imagem inicial 2 $(ET=1.1520,\, EF=8.4903\times 10^{-4})$



Figura 7.7: Após 100 iterações da versão 2 do algoritmo IOH, considerando a imagem inicial 2 ($ET = 1.1449, EF = 1.02110 \times 10^{-3}$)

Exemplo 2



Figura 7.8: Após 100 iterações do algoritmo ER, considerando a imagem inicial 1 (ET = 1.07457, $E0 = 1.04832 \times 10^{-1}$, $EF = 1.04816 \times 10^{-1}$)



Figura 7.9: Nova imagem inicial gerada pela reflexão através do centro do toróide



Figura 7.10: Após 100 iterações do algoritmo ER (ET = 1.03290, E0 = 0.182176, EF = 0.182170)



Figura 7.11: Nova imagem inicial gerada pela reflexão através do centro do toróide



Figura 7.12: Após 100 iterações do algoritmo ER (
 $ET = 1.045525, E0 = 6.48480 \times 10^{-2},$
 $EF = 6.48438 \times 10^{-2})$

Exemplo 3



Figura 7.13: Após 100 iterações do algoritmo ER, utilizando a imagem inicial 2 (ET = 0.9013, $E0 = 1.180862 \times 10^{-1}$, $EF = 1.180058 \times 10^{-1}$)



Figura 7.14: Nova imagem inicial gerada pela reflexão através do centro do toróide



Figura 7.15: Após 100 iterações do algoritmo ER (ET = 1.1681, E0 = 0.18114, EF = 0.18093)



Figura 7.16: Nova imagem inicial gerada pela reflexão através do centro do toróide



Figura 7.17: Após 100 iterações do algoritmo ER (
 $ET=0.88375561,\ E0=7.058566\times 10^{-2},\ EF=7.061454\times 10^{-2})$
Exemplo 4



Figura 7.18: Após 200 iterações do algoritmo ER utilizando a imagem inicial 3 (ET = 1.348260, E0 = 0.1161011, EF = 0.116089)



Figura 7.19: Nova imagem inicial gerada pela reflexão através do centro do toróide



Figura 7.20: Após 500 iterações do algoritm
o $\mathrm{ER}(ET=1.228234,\,E0=9.754321\times10^{-2},\,EF=9.754321\times10^{-2})$

Exemplo 5



Figura 7.21: Após 100 iterações do algoritmo IOH utilizando a imagem inicial 1 ($ET=1.14730,\, EF=8.9714\times 10^{-4})$



Figura 7.22: Nova imagem inicial gerada pela reflexão através do centro do toróide



Figura 7.23: Após 100 iterações do algoritmo IOH ($ET=0.6237,\, EF=1.01442\times 10^{-3}$



Figura 7.24: Nova imagem inicial gerada pela reflexão através do centro do toróide



Figura 7.25: Após 100 iterações do algoritmo IOH ($ET=1.1937,\, EF=1.8232\times 10^{-3})$



Figura 7.26: Nova imagem inicial gerada pela reflexão através do centro do toróide



Figura 7.27: Após 100 iterações do algoritmo IOH ($ET=0.9135,\, EF=6.0248\times 10^{-4})$

Capítulo 8

CONCLUSÕES

Nesta dissertação introduzimos uma nova forma de analisar o problema da recuperação da fase, no domínio do espaço, através da geometria do problema discretizado em dimensão finita. Esta nova forma de ver o problema originou uma prova da limitação de alguns dos algoritmos assim como um novo método automático para sair dos pontos de estagnação. Introduzimos também algumas modificações em métodos iterativos, já existentes, para resolver o problema (Cap. 3).

A recuperação da fase é um problema difícil, não somente porque os algoritmos estagnam, mas também porque na prática é muitas vezes difícil decidir se o algoritmo encontrou uma solução, estagnou porque o algoritmo entrou num ciclo, ou estagnou num ponto que satisfaz os dados com uma precisão muito grande, sem ser solução. Até agora a forma de enfrentar o problema tem sido, do ponto de vista prático, através da análise do tipo de imagens representadas pelos pontos de estagnação, e, do ponto de vista teórico, considerando a fatoração da função no domínio de Fourier. A nossa aproximação, com a caracterização geométrica das funções no domínio do espaço, gera uma nova possibilidade de dar uma resposta aos inconvenientes mencionados. A quantidade de perguntas que surgiram, e continuam surgindo por causa deste novo ponto de vista, e das experiências numéricas induzidas por ele, é enorme. Enumeramos apenas algumas, sem que a ordem esteja relacionada com a importância:

1) Por que os pontos fixos dos algoritmos envolvem sempre apenas um toróide e as restrições convexas, e não outro toróide?

2) Quais são as características das imagens contidas em cada um dos toróides?

3) Qual é o comportamento dos toróides quando os dados tendem a um contínuo, já que sabemos que no caso contínuo tem unicidade com probabilidade um?

4) Qual seria a descrição geométrica do caso sem unicidade?

5) É possível introduzir algum tipo de regularização no problema que evite a estagnação ou que pelo menos a dificulte?

6) Existem outras alternativas melhores para a reflexão sugerida como método no Capítulo6?

Continuaremos trabalhando nas direções sugeridas pelas perguntas acima e por muitas outras. Também nos propomos a trabalhar na aplicação das mesmas idéias aos outros problemas relacionados, descritos na Introdução.

Bibliografia

- Akutowics, E., "On the determination of the phase of the Fourier integral, I, II", Trans. Amer. Math.Soc., 83(1956), 179-192, 84(1957), 237-238.
- [2] Ayers, G.R. and Dainty, J.C., "An iterative blind deconvolution method and its applications", Opt. Lett. 13, 547-549, 1988.
- [3] Barakat, R. and Newsam, G., "Algorithms for reconstruction of partially know, bandlimited Fourier-transform pairs from noisy data", J. Opt. Soc. Am. A, vol 2, 11, 1985.
- [4] Bates, R.H.T., "Astronomical speckle imaging", Physics Reports, 90, 203, 1982.
- [5] Bates, R.H.T. and McDonnell, M.J., "Image Restoration and Reconstruction", Oxford Science Publications, 1986.
- [6] Bracewell, R., "The Fourier Transform and its Applications", New York, Mc. Grawhill, 1965.
- [7] Brigham, E.O., "The Fast Fourier Transform and its Applications", Prentice-hall, Inc, 1988.
- [8] Bruck, M. and Sodin, L.G., "On the ambiguity of the Image Reconstruction Problem", Optics Communications, vol. 30, 3, 304-308, 1979.
- [9] Chen, P. and Fiddy, M., "Image reconstruction from power spectral data using point zero locations", Journal of the Optical Society of America, A, 11, 2210-2214, 1994.

- [10] Combettes, P.L. and Trussell, H.J., "Method of Successive Projections for Finding a Common Point of Sets in Metric Spaces", Journal of Otimization Theory and Applications, vol 67, n^Q 3, 1990.
- [11] Crimmins, T.R., Fienup, J.R. and Thelen, B.J., "Improved bounds on object support from autocorrelation support and application to phase retrieval", J. Opt. Soc. Am. A, vol 7, 1, 1990.
- [12] Dainty, J.C. and Fienup, J.R., "Phase retrieval and image reconstruction for astronomy", in: "Image Recovery: Theory and Application", H. Stark, Ed), Academic Press, Inc., 1987.
- [13] Davey, B.L., Lane, R.G. and BATES, R.H.T., "Blind deconvolution of noisy complexvalued image", Opt. Commun 69, 353-356, 1989.
- [14] Dianat, S.A. and Raghuveer, M.R., "Fast algorithms for phase and magnitude reconstruction from bispectra", Optical Engineering, vol 29, n^Q 5, 504-512, 1990.
- [15] Fienup, J.R., "Improved Synthesis and Computational Methods for Computer-Generated Holograms", PHd Dissertation, Stanford University, 1975.
- [16] Fienup, J.R., "Reconstruction of an object from the modulus of its Fourier transform", Optics Letters, vol 3, 1, 1978.
- [17] Fienup, J.R., "Phase Retrieval Algorithms: A comparison", Applied Optics, vol 21, 15, 2758-2769, 1982.
- [18] Fienup, J.R., Crimmins, T.R., and W. Holsztynski, W., "Reconstruction of the support of an object from the support of its autocorrelation", J. Opt. Soc. Am. 72, 610-624, 1982.
- [19] J. R. Fienup, J.R. and Wackerman, C., "Phase-retrieval stagnation: problems and solutions", J. Opt. Soc. Am. A, vol 3, 11, 1897-1907, 1986.

- [20] Gerchberg, R.W. and Saxton, W.O., "A pratical algorithm for the determination of phase from image and diffraction plane pictures", Optik, vol 35, 2, 237-246, 1972.
- [21] Labeyre, A., "Attainment of diffraction limited resolution in large telescopes by Fourier analysis speckle patterns in stars images", Astronomy and Astrophysics, 6, 85-, 1970.
- [22] Hayes, M.H., "The reconstruction of a multidimensional sequence from the phase or magnitude of its Fourier transform", IEEE Trans. Acoust. Speech Processing, ASSP-30, 140-155, 1982.
- [23] M. H. Hayes, M.H., Lim, J.S. and Oppenheim, A.V., "Signal reconstruction from phase or magnitude", IEE ASSP-28, 672-680, 1980.
- [24] Herman, G.T., and Ro D.W., "Image recovery using iterative data refinement with relaxation", Medical Image Processing Group, Departement of Radiology, University of Pennsylvania, Technical Report no.MIPG156, 1989.
- [25] Hofstetter, E., "Construction of time-limited functions with specified autocorrelation functions", IEEE Trans. Inform. Theory, IT-10, 119-126, 1964.
- [26] Hurt, N.E., "Phase Retrieval and Zero Crossings: Mathematical Methods in Image Reconstruction", Kluwer Academic Publishers, 1989.
- [27] Lane, R.G., "Phase retrieval using conjugate gradient minimization", Journal of Modern Optics, vol 38, 9, 1797-1813, 1991.
- [28] Lane, R.G., "Blind deconvolution of speckle images", J. Opt. Soc. Am. A, vol 9, 9, 1992.
- [29] Levi, A. and Stark, H., "Image restoration by the method of generalized projections whith application the restoration from magnitude", J. Opt. Soc. Am. A, vol1, 9, 1984.

- [30] Meyn, K., "Solution of underdetermined nonlinear equations by stationary iteration methods", Numer. Math.42, 161-172, 1982.
- [31] Millane, R.P., "Phase retrieval in cristallography and optics", J. Opt. Soc. Am. A, vol 7, 3, 1990.
- [32] Misell, D.L., "The phase problem in electron microscopy", in Advances in Optical and Electron Microscopy (eds. R. Barer and V.E. Colett),7, 185, 1978.
- [33] Mou-yan, Z. and Unbehauen, R., "Methods for reconstruction of 2-D sequences from Fourier transform magnitude", IEEE Transactions on Image Processing, 6, 2, 222-233, 1997.
- [34] Nakajima, N., "Blind deconvolution of a Hermitian and a non-Hermitian function", J. Opt. Soc. Am. A, vol 8, 5, 1991.
- [35] Napier, P.J. and Bates, R.H.T., "Inferring phase information from modulus information in two-dimensional aperture synthesis", Astron. Astrophys. Suppl. 15, 427-430, 1974.
- [36] O'Neill, E.L. and Walter, A., "The question of phase in image formation", Opt. Acta 10, 33-40, 1963.
- [37] M. Nieto-Vesperinas, M., "A study of the performance of nonlinear least-square optimization methods in the problem of phase retrieval", Optica Acta, vol 33, 6, 713-722, 1986.
- [38] M. Nieto-Vesperinas, M., and Mendes, J.A., "Phase retrieval by Monte Carlo methods", Opt. Commun. 59, 249-254, 1986.
- [39] Noushin, A.J., Fiddy, M.A. and Pommet, D.A., "Optical imaging from scattering data: Imaging from fourier intensity data", preprint, Dept. Electrical Engineering, University of Massachussetts, Lowell, 1997.

- [40] Oppenheim, A.V. and Schaffer, R.W., "Digital Signal Processing", Prentice Hall, Englewood Cliffs, N.J., 1975.
- [41] Ortega, J.and Rheinboldt, W. C., "Iterative Solution of Nonlinear Equations in Several Variables", New York: Academic Press, 1970.
- [42] A. Papoulis, "The Fourier Integral and its Applications", New York, Mc Graw-Hill, 1962.
- [43] Pierra, G., "Decomposition Through Formalization in a Product Space", Mathematical Programming 28, 96-115, 1984.
- [44] Ramchandran, G.N. and Srinivasan, R., "Fourier Methods in Crystallografy", Wiley, New York, 1970.
- [45] Sanz, J.L.C., "Mathematical considerations for the Problem of Fourier transform Phase Retrieval from magnitude", SIAM J. Appl. Math., vol 45, 4, 651-664, 1985.
- [46] SANZ, J.L. and Huang, T.S., "Stability of unique Fourier transform phase reconstruction", J. Opt. Soc. Am., vol 73, 11, 1442-1445, 1983.
- [47] Sayre, D., "The squaring method: a new method for phase determination", Acta Crystallogr., 5, 60-65, 1952.
- [48] Seldin, J.H. and Fienup, J.R., "Numerical investigation of the uniquenes of phase retrieval", Optical Society of America A, vol 7, 3, 412-427, 1990.
- [49] Seldin, J.H. and Fienup, J.R., "Iterative blind deconvolution algorithm applied to phase retrieval", Optical Society of America A, vol 7, 3, 428-433, 1990.
- [50] Walter, A., "The question of phase retrieval in optics", Opt. Acta, 10 (1963), 41-49.
- [51] Yellott Jr., J.I. and GEOFFREY J. Iverson, G.J. "Uniques Properties of Higher-order Autocorrelation Functions" J. Opt. Soc. Am. vol. 9, 3, 388-404,1992.

[52] Youla, D.C., "Mathematical theory of image restoration by the method of convex projections", in Image Recovery: Theory and Applications, H. (Stark, ed.), Academic Press, Orlando, Fla, 29-77, 1987.