"METODOLOGIA PARA ESTIMAÇÃO DAS CONSTANTES ÓPTICAS E DA ESPESSURA DE UM FILME FINO UTILIZANDO MINIMIZAÇÃO IRRESTRITA".

Refere-se o presente relatório a uma metodologia para destinada a possibilitar a estimação das constantes ópticas e da espessura de um filme fino, que caracteriza-se pelo fato de estar baseada na minimização irrestrita.

O problema de estimar a espessura e as constantes óticas de um filme fino utilizando só dados de transmissão é constitui uma tarefa muito desafiante do ponto de vista matemático e tem importância tecnológica e econômica. Em muitos casos um problema inverso muito mal condicionado com muitas soluções locais não-globais. Numa publicação recente foi proposto pelos titulares da presente patente, modelos de programação não linear para resolver este problema. Bem conhecidos softwares para minimização com restrições lineares foram utilizados com sucesso. Neste trabalho introduziu-se uma formulação irrestrita do modelo de programação não linear e resolvendo assim o problema de estimação utilizando um método baseado em chamadas sucessivas a um algoritmo de minimização irrestrita. Experimentos numéricos em filmes gerados no computador mostram que o novo procedimento é eficiente.

O problema consiste em achar a espessura d, o índice de refração $r(\lambda)$ e o coeficiente de atenuação $k(\lambda)$ de um filme fino, utilizando só dados de transmissão. Emite-se luz com diferentes comprimentos de onda λ_i e mede-se a luz transmitida T_i^{meas} do outro lado do filme. Com um conjunto razoável de medições (λ_i, T_i^{meas}) tentou-se descobrir as constantes óticas acima mencionadas. A espessura dos filmes é um importante parâmetro de desenho e caracterização. A transmissão ótica fornece informação precisa no intervalo do espectro onde o material vai da opacidade completa a algum grau de transparência (M. BORN e E. WOLF 1980; e O.S.HEAVENS,1991).

Algumas soluções aproximadas têm sido encontradas em casos onde a transmissão mostra um padrão de interferência numa região altamente transparente do espectro (J.C. MANIFACIER, J. GASIOT, e J.P. FILLARD 1976; R. SWANEPOEL 1983; e R. SWANEPOEL 1984).

Porém, até agora, a solução geral do problema não tem sido satisfatória porque o sistema de equações é altamente indeterminado.

Recentemente, foi reportado um método novo, baseado num enfoque de minimização "ponto a ponto" com restrições, que permite resolver o caso geral (I.CHAMBOULEYRON, J.M.MARTÍNEZ, A.C. MORETTI e M.MULATO 1988; e I.CHAMBOULEYRON, J.M.MARTÍNEZ, A.C. MORETTI e M.MULATO 1997).

O método define um problema de programação não-linear, cujas variáveis são os coeficientes a serem estimados, com restrições que representam conhecimentos prévios da solução física.

O novo método foi bem sucedido na recuperação de d, $r(\lambda)$ e $k(\lambda)$ a partir de diferentes espectros de transmissão de filmes artificiais e reais(I.CHAMBOULEYRON, J.M.MARTÍNEZ, A.C. MORETTI e M.MULATO 1988; e I.CHAMBOULEYRON, J.M.MARTÍNEZ, A.C. MORETTI e M.MULATO 1997).

A maior dificuldade do enfoque do problema de otimização ponto a ponto com restrições é que um problema de grande porte relativamente complexo de programação não-linear com restrições lineares cuja solução só pode ser obtida por meio de códigos sofisticados, e não sempre disponíveis, que trabalhem eficientemente com a esparsidade da matriz de restrições (B.A.MURTAGH e M.A.SAUNDERS 1977; e B.A.MURTAGH e A.SAUNDERS 1978).

Considerou-se o problema de estimar o coeficiente de atenuação, o índice de refração e a espessura do filme, utilizando só dados de transmissão. Dado o comprimento de onda λ, o índice de refração do substrato s, e as incógnitas d (espessura), $r(\lambda)$(índice de refração) e $k(\lambda)$ (coeficiente de atenuação), a transmissão teórica é dada por uma fórmula conhecida (O.S.HEAVENS 1991; R.SWANEPOEL 1983; e R.SWANEPOEL 1984).

Tendo medidas da transmissão em (bastantes) comprimentos de onda diferentes, buscou-se estimar as incógnitas acima mencionadas. Numa primeira avaliação, verifica-se que este problema é altamente indeterminado. Para cada cumprimento de onda, a equação

Transmissão teórica = Transmissão medida (1)

tem três incógnitas d, $r(\lambda)$, $k(\lambda)$ e só d repete-se para todos os valores de λ. A idéia principal proposta por I.

CHAMBOULEYRON em 1998 e 1997 foi incorporar conhecimentos prévios das funções para diminuir o grau de liberdade de (1) ao ponto que só estimativas com significado físico sejam aceitas.

A idéia de supor fórmulas fechadas para r e k dependendo de uns poucos parâmetros tem sido reportada em por O.S.HEAVENS em 1991, R.SWANEPOEL em 1983 e 1984.

Os métodos originados nessas idéias são eficientes quando a curva de transmissão tem um padrão de interferência em regiões relativamente grandes do espectro onde $k(\lambda)$ é quase nula. Em outros casos, a satisfação de (1) é muito grosseira ou as curvas $r(\lambda)$ e $k(\lambda)$ são fisicamente inaceitáveis.

I.CHAMBOULEYRON, J.M. MARTÍNEZ, A.C.MORETTI e M.MULATO, em 1997 e 1998, defendem que no lugar de impor formas funcionais para $r(\lambda)$ $k(\lambda)$, as restrições fenomenológicas que restringem a variabilidade dessas funções foram consideradas explicitamente de forma tal que o problema de estimação tomou a forma:

Minimizar $\underset{\lambda}{K}$ [Transmissão teórica (λ) - Transmissão medida (λ)]^2 sujeito a Restrições Físicas. (2)

Desta forma, funções $r(\lambda)$ e $k(\lambda)$ bem comportadas podem ser obtidas sem restrições fortes que poderiam danificar a qualidade do ajuste (1).

A maior contribuição deste trabalho é estabelecer um método para resolver o problema de estimação onde (2) é substituído por um problema de otimização sem restrições.

Resolveu-se este problema utilizando o Método do Gradiente Espectral (SGM) introduzido recentemente por M.Raydan em 1997. Este método baseia-se numa idéia muito efetiva para problemas de minimização irrestrita de grande porte e consiste em utilizar só direções de gradiente com comprimentos de passo que garantem uma convergência rápida. A redução de (2) a um problema de minimização sem restrições precisou do cálculo de derivadas complicadas que não teria sido possível sem a utilização de técnicas de diferenciação automática. Aqui utilizamos o modo reverso de diferenciação automática descrito por E.G.BIRGIN em 1998.

Na formulação irrestrita do problema de estimação, a transmissão T de um filme fino absorvente depositado num substrato transparente grosso (R.SWANEPOEL 1983 E 1984)é dado por:

$$\mathrm{T} = \frac{Ax}{B - Cx + Dx^{2}}, \tag{3}$$

onde

$$A = 16s(r^2 + k^2) \tag{4}$$

$$B = [(r+1)^2 + k^2)][(r+1)(r+s^2) + k^2\}, \tag{5}$$

$$C = [(r^2 - 1 + k^2)(r^2 - s^2 + k^2) - 2k^2(s^2+1)]2\cos\varphi$$

$$-k[2(r^2 - s^2 + k^2) + (s^2+1)(r^2 - 1 + k^2)]2\,\mathrm{sen}\,\varphi\,, \tag{6}$$

$$D = [(r-1)^2 + k^2][(r-1)(r-s^2) + k^2], \tag{7}$$

$$\varphi = 4\pi rd/\lambda,\ x = \exp(-\alpha d),\ \alpha = 4\pi k/\lambda. \tag{8}$$

Nas fórmulas (4)-(8) é utilizada a seguinte notação:

(a) λ é o comprimento de onda;

(b) $s = s(\lambda)$ é o índice de refração do substrato transparente (que supomos conhecido);

(c) $r = r(\lambda)$ é o índice de refração do filme;

(d) $k = k(\lambda)$ é o índice de atenuação do filme (α é o coeficiente de absorção);

(e) d é a espessura do filme.

Dado um conjunto de dados experimentais $(\lambda_i, T^{meas}(\lambda_i))$, $\lambda_{min} \le \lambda_i < \lambda_{i+1} \le \lambda_{max}$, para $i = 1,\ldots,N$, queremos estimar d, $r(\lambda)$ e $\kappa(\lambda)$. Este problema parece altamente indeterminado. De fato, para d conhecido e um λ dado, a seguinte equação deve ser satisfeita:

$$T(\lambda, s(\lambda), d, r(\lambda), k(\lambda)) = T^{meas}(\lambda). \qquad (9)$$

Esta equação tem duas incógnitas $r(\lambda)$ e $k(\lambda)$ e, portanto, em geral, o seu conjunto de soluções é uma curva no espaço bidimensional $(r(\lambda), k(\lambda))$. Portanto, o conjunto de funções (r,k) que satisfazem (9) para um dado d é infinito e, informalmente, é representado por uma variedade não-linear de dimensão N em R^{2N}.

Porém, restrições físicas reduzem drasticamente o conjunto viável das incógnitas $r(\lambda)$ e $k(\lambda)$. Por exemplo, na vizinhança do eixo fundamental de absorção (dispersão normal), essas restrições físicas são:

PC1: $r(\lambda) \ge 1$, $k(\lambda) \ge 0$ para todo $\lambda \in [\lambda_{min}, \lambda_{max}]$;

PC2: $r(\lambda)$ e $k(\lambda)$ são funções decrescentes de λ;

PC3: $r(\lambda)$ é convexa;

PC4: existe $\lambda_{infl} \in [\lambda_{min}, \lambda_{max}]$ tal que $k(\lambda)$ é convexa se $\lambda \ge \lambda_{infl}$ e côncava se $\lambda < \lambda_{infl}$.

Observe-se que PC1 é satisfeita supondo PC2 válida, $r(\lambda_{max}) \geq 1$ e $k(\lambda_{max}) \geq 0$. As restrições PC2, PC3 e PC4 podem ser escritas, respectivamente, como

$n'(\lambda) \leq 0$ e $k'(\lambda) \leq 0$ para todo $\lambda \in [\lambda_{min}, \lambda_{max}]$; (10)

$n''(\lambda) \geq 0$ para todo $\lambda \in [\lambda_{min}, \lambda_{max}]$; (11)

$k''(\lambda) \leq 0$ para todo $\lambda \in [\lambda_{min}, \lambda_{\inf l}]$ (12)

$k''(\lambda) \geq 0$ para todo $\lambda \in [\lambda_{\inf l}, \lambda_{max}]$ (13)

Claramente, as restrições

$n''(\lambda) \geq 0$ para todo $\lambda \in [\lambda_{min}, \lambda_{max}]$ e $n'(\lambda_{max}) \leq 0$

implicam

$n'(\lambda) \leq 0$ para todo $\lambda \in [\lambda_{min}, \lambda_{max}]$

Mais ainda,

$k''(\lambda) \geq 0$ para todo $\lambda \in [\lambda_{\inf l}, \lambda_{max}]$ e $k'(\lambda_{max}) \leq 0$

implicam

$k'(\lambda) \leq 0$ para todo $\lambda \in [\lambda_{\inf l}, \lambda_{max}]$

Finalmente,

$k''(\lambda) \leq 0$ para todo $\lambda \in [\lambda_{min}, \lambda_{\inf l}]$ e $k'(\lambda_{min}) \leq 0$

implicam

$k'(\lambda) \leq 0$ para todo $\lambda \in [\lambda_{min}, \lambda_{\inf l}]$

Portanto, PC2 pode ser substituída por

$n'(\lambda_{max}) \leq 0,\ k'(\lambda_{max}) \leq 0,\ K'(\lambda_{min}) \leq 0.$ (14)

Resumindo, as restrições PC1-PC4 serão satisfeitas se, e somente se,

$r(\lambda_{max}) \geq 1,\ k(\lambda_{max}) \geq 0,$ (15)

$n'(\lambda_{max}) \leq 0,\ k'(\lambda_{max}) \leq 0,$ (16)

$n''(\lambda) \geq 0$ para todo $\lambda \in [\lambda_{min}, \lambda_{max}],$ (17)

$k''(\lambda) \geq 0$ para todo $\lambda \in [\lambda_{\inf l}, \lambda_{max}],$ (18)

$k''(\lambda) \leq 0$ para todo $\lambda \in [\lambda_{min}, \lambda_{\inf l}]$ (19)

$$k'(\lambda_{min}) \leq 0 \quad (20)$$

Logo, a solução de quadrados mínimos contínuos do problema de estimação é a solução $(d, r(\lambda), k(\lambda))$ de

$$\text{Minimizar} \int_{\lambda min}^{\lambda max} | T(\lambda, s(\lambda), d, r(\lambda), k(\lambda)) - T^{meas}(\lambda)^2 d\lambda \quad (21)$$

sujeito às restrições (15)-(20).

O objetivo desta patente é eliminar as restrições do problema, tanto quanto seja possível, por meio de mudança de variáveis. Informalmente, será escrita a função objetivo (21) dependendo das derivadas segundas de $r(\lambda)$ e $k(\lambda)$ e os valores das funções e suas derivadas primeiras em λ_{max}. Além disso, a possitividade será garantida expressando as variáveis como quadrados de variáveis auxiliares. De fato, escreveu-se

$$r(\lambda_{max}) = 1 + u^2,\ k(\lambda_{max}) = v^2 \quad (22)$$

$$n'(\lambda_{max}) = -u_1^2,\ k'(\lambda_{max}) = -v_1^2 \quad (23)$$

$$n''(\lambda) = w(\lambda)^2 \text{ para todo } \lambda \in [\lambda_{min}, \lambda_{max}], \quad (24)$$

$$k''(\lambda) = z(\lambda)^2 \text{ para todo } \lambda \in [\lambda_{infl}, \lambda_{max}] \quad (25)$$

$$k''(\lambda) = -z(\lambda)^2 \text{ para todo } \lambda \in [\lambda_{min}, \lambda_{infl}] \quad (26)$$

Neste ponto, para evitar uma enfadonha formulação contínua do problema, considera-se a situação real, na qual os dados são representados por um conjunto de N pontos igualmente espaçados no intervalo $[\lambda_{min}, \lambda_{max}]$. Portanto, defini-se

$$h = (\lambda_{max}, \lambda_{min}) / (N-1).$$

$$\lambda_1 = \lambda_{min} + (i-1)h,\ i = 1, \ldots, N$$

A transmissão medida em λ_i, será chamada T_i^{meas}. Será usada a notação r_i, k_i, w_i, z_i para os estimadores por diferenças finitas de $r(\lambda_i), k(\lambda_i), w(\lambda_i)$ e

$z(\lambda_i)$:

$r_i \approx r(\lambda_i)$, $\qquad$ $k_i \approx k(\lambda_i)$,

$w_i \approx w(\lambda_{i+1})$, $\qquad$ $z_i \approx z(\lambda_{i+1})$,

para $i = 1,\ldots,N$. A discretização das relações diferenciais (22-26) leva a:

$$r_N = 1 + u^2, u_N = v^2, \tag{27}$$

$$r_{N-1} = r_N + u_1^2 h, \quad k_{N-1} = k_N + v_1^2 h, \tag{28}$$

$$r_i = w_1^2 h^2 + 2n_{i+1} - r_{1+2}, \quad i = 1,\ldots,N-2 \tag{29}$$

$$k_i = z_1^2 h^2 + 2k_{i+2} - k_{i+2}, \quad \text{se} \quad \lambda_{i+1} \geq \lambda_{\mathrm{inf}l}, \tag{30}$$

$$k_i = -z_1^2 h^2 + 2k_{i+2} - k_{i+2}, \quad \text{se} \quad \lambda_{i+1} < \lambda_{\mathrm{inf}l}, \tag{31}$$

Finalmente a função objetivo (21) é aproximada por uma soma de quadrados, levando ao problema de otimização

$$\text{Minimizar} \quad \sum_{i=1}^{N} [T(\lambda_i, s(\lambda_i), d, r_i, k_i) - T_i^{meas}]^2 \tag{32}$$

sujeito a

$$k_1 \geq k_2 \tag{33}$$

Como r_i e k_i dependem de u, u_i, v, v_i, w, z e $\lambda_{\mathrm{inf}l}$ por (27-31), problema (32) toma a forma

Minimizar $f(d, \lambda_{\mathrm{inf}l}, u, u_1, v, v_1, w_1, \ldots, w_{N-2}, z_1, \ldots, z_{N-2}$ sujeito a (33).

Como espera-se que a restrição (33) não seja ativa na solução de (34-33), será considerado o problema irrestrito (34). As variáveis presentes em (34) têm natureza diferente. A espessura d é uma variável dimensional (medida em nanômetros nos problemas reais) que pode ser determinada utilizando as observações $T^{meas}(\lambda_i)$ para $\lambda_i \geq \lambda_{bound}$, onde λ_{bound}, um limitante superior de $\lambda_{\mathrm{inf}l}$, reflete nosso conhecimento prévio do problema. Por essa razão, o primeiro passo no processo de estimação será estimar

d com os dados que correspondem a $\lambda_i \geq \lambda_{bound}$. Para atingir este objetivo resolveu-se o problema

$$\text{Minimizar } \bar{f}(u,u_1,v,v_1,w,z) \equiv \sum_{\lambda_i \geq \lambda_{bound}} [T(\lambda_i, s(\lambda_i), d, r_i, k_i) - T_i^{meas}]^2 \qquad (35)$$

para diferentes valores de d e tomou-se como espessura estimada aquela que corresponda ao menor valor de função. Neste caso, a restrição (33) é irrelevante porque é automaticamente satisfeita pela convexidade de k e o fato de que a derivada primeira de k em λ_{max} é não-positiva. Daqui em diante será considerado que d está fixa como resultado do procedimento anterior.

O segundo passo consiste em determinar λ_{infl}, junto com u,u_1,v,v_1,w,z. Com este propósito, observe-se que, dados d e λ_{infl} o problema

$$\text{Minimizar} \sum_{i=1}^{N} [T(\lambda_i, s(\lambda_i), d, r_i, k_i) - T_i^{meas}]^2 \qquad (36)$$

é (desprezando (33)) um problema de minimização irrestrita cujas variáveis são u,u_1,v,v_1,w,z, (2N variáveis). Resolveu-se esse problema para vários valores de λ_{infl} e tomou-se como estimativas de r e k a combinação de variáveis correspondente ao menor valor da função objetivo. Para minimizar esta função e resolver (35) para diferentes espessuras, utilizou-se o algoritmo de minimização irrestrita descrito a seguir.

Como visto acima, os problemas de minimização irrestrita (35) e (36) têm a forma

$$\text{Minimizar } f(u,u_1,v,v_1,w_1,\ldots w_{N-2},z_1,\ldots,z_{N-2}). \qquad (37)$$

Para simplificar a notação, nesta seção será escrito $x = (u,u_1,v,v_1,w_1,\ldots,w_{N-2},z_1,\ldots,z_{N-2})$.

As derivadas parciais de f

são usualmente necessárias em algoritmos de otimização, porque fornecem a informação de primeira ordem da função objetivo que permite seguir trajetórias de descida. Neste caso, as derivadas são difíceis de calcular e por esta razão utilizou-se o modo reverso de diferenciação automática para o cálculo das derivadas.

Em princípio, qualquer algoritmo de otimização irrestrita pode ser utilizado para resolver (37).

Como o problema tem, potencialmente, um número grande de variáveis, a escolha deve se restringir aos métodos que sejam capazes de encarar essa situação. Um trabalho recente de M.RAYDAN(1997)nos levou a escolher o Método do Gradiente Espectral (SGM), uma implementação do método de Barzilai and Borwein para quadráticas introduzido no trabalho de M.RAYDAN (1997). De fato, Raydan mostrou, utilizando um conhecido conjunto de problemas teste clássicos, que SGM superou métodos de gradientes conjugados em problemas de otimização irrestrita de grande porte. O método do gradiente espectral de Raydan é fácil de implementar, um fato que contribuiu em nossa decisão, porque permite-nos ficar independentes de, por exemplo, softwares fechados. A descrição do SGM é essencialmente, a mesma apresentada por M.RAYDAN com uma pequena diferença na escolha do passo α_k quando $b_k \leq 0$.

Denotamos $g(x) = \nabla f(x)$. O algoritmo começa com $x_0 \in R^n$ e utiliza um inteiro $M \geq 0$, um parâmetro pequeno α_{min}, um parâmetro grande $\alpha_{max} > \alpha_{min}$ um parâmetro de decréscimo suficiente $\gamma \in (0,1)$ e parâmetros de salvaguarda $0<\sigma_1,<\sigma_2,<1$. Inicialmente, $\alpha_0 \in [\alpha_{min}, \alpha_{max}]$ é arbitrário. Dado $x_k \in$ R^n e $\alpha_k \in [\alpha_{min}, \alpha_{max}]$, o Algoritmo 3.1 descreve como obter x_{k+1} e α_{k+1}, e quando terminar o processo.

Será feito agora a descrição do algoritmo 3.1, o qual compreende: Passo 1: Detectar se o ponto atual é estacionário. Se $\|g(x_k)\| = 0$, parar declarando que x_k é estacionário.

Passo 2: Busca linear.

Passo 2.1: Fazer $\lambda \leftarrow \alpha_k$.

Passo 2.2: Fazer $x_+ = x_k - \lambda g(x_k)$.

Passo 2.3: Se $f(x_+) \leq max_{0\leq j \leq min\{k,M-1\}} f(x_{k-j}) + \gamma \langle x_+ - x_{k,} g(x_k) \rangle$, (38)

então definir $x_{k+1} = x_+, s_k = x_{k+1} - x_k e y_k = g(x_{k+1}) - g(x_k)$.

Senão, definir $\lambda_{new} \in [\sigma_1 \lambda, \sigma_2 \lambda]$, (39)

Fazer $\lambda \leftarrow \lambda_{new}$ e voltar ao Passo 2.2.

Passo 3: Calcular o comprimento de passo espectral.

Calcular $b_k = \langle s_k, y_k \rangle$

Se $b_k \leq 0$, fazer $\alpha_{k+1} = \alpha_{max}$

senão, calcular $\alpha_k = -\langle s_k, g(x_k) \rangle$ e

$\alpha_{k+1} = \min\{\alpha_{max}, \max\{\alpha_{min}, a_k / b_k\}\}$

Na prática, o cálculo de λ_{new} utiliza uma interpolação quadrática unidimensional com as salvaguardas (39).

Os resultados numéricos serão agora descritos.

Para testar a confiabilidade do novo enfoque de minimização irrestrita, utilizamos filmes artificiais (criados no computador) depositados em substratos de vidro ou silício cristalino. Nas simulações, o índice de refração do vidro $s_{glass}(\lambda)$ é dado por:

$s_{glass}(\lambda)=\sqrt{1+1/(0.76194-740/\lambda^2)}$,

e o índice de refração do substrato de silício $s_{Si}(\lambda)$ é dado por: $s_{Si}(\lambda)=3.71382-8.6912310^{-5}\lambda-2.4712510^{-8}\lambda^2+1.0467710^{-11}\lambda^3$

Em todas as simulações supomos que o comprimento de onda e a espessura são medidos em nanômetros. A transmissão $T^{meas}(\lambda)$ para cada filme foi gerada no intervalo $\lambda\in[\lambda_{min},\lambda_{max}]$ utilizando espessura d_{true}, índice de refração $r_{true}(\lambda)$ e coeficiente de absorção $\alpha_{true}(\lambda)$? conhecidos. Para considerar situações realísticas, incluindo a falta de precisão experimental, considerou-se cálculos alternativos de $T^{meas}(\lambda)$, onde a transmissão real (gerada) foi arredondada para 4, 3 e 2 casas decimais. Também foram realizados experimentos numéricos utilizando diferentes números de pontos de transmissão: 100, 50 e 25. A descrição dos filmes gerados e os correspondentes resultados numéricos são apresentados abaixo:

Filme A: Este filme simula um filme fino de germânio amorfo com d_{true} = 118nm depositado num substrato de vidro $\lambda_{min}=600\text{nm}$ e $\lambda_{max}=2000\text{nm}$.

Filme B: Este filme é idêntico ao Filme A com d_{true}=782nm. $\lambda_{min}=1000\text{nm}$ e $\lambda_{max}=2000\text{nm}$.

Filme C: Este filme simula um filme fino de germânio amorfo com d_{true}=147nm depositado em um substrato cristalino de silício. $\lambda_{min}=1250\text{nm}$ e $\lambda_{max}=2500\text{nm}$.

Filme D: Este filme é idêntico ao Filme C com d_{true}=640nm. $\lambda_{min}=640\text{nm}$ e $\lambda_{max}=640\text{nm}$.

Filme E: Este filme simula um filme fino de silício amorfo hidrogenado com d_{true}=624nm depositado em vidro. $\lambda_{min} = 600\text{nm}$ e $\lambda_{max} = 1600\text{nm}$.

Para os nossos cálculos, necessitamos estimativas iniciais de $k(\lambda)$ e $r(\lambda)$. Como estimativa inicial de $k(\lambda)$ utilizou-se uma função linear por partes cujos valores são 0.1 em λ_{min}, 0.01 em $+ 0.2\,(\lambda_{max} - \lambda_{min})$ e 10^{-10} em λ_{max} . A estimativa inicial de $r(\lambda)$ foi uma função linear cujos valores extremos variam entre 5 e 3 com passo 1 (esses valores foram escolhidos pelo nosso conhecimento prévio dos materiais simulados). Como excluímos as funções constantes que mostraram em testes preliminares levar o método a mínimos locais, temos três possibilidades para a estimativa inicial de $r(\lambda)$: as funções lineares decrescentes definidas pelos pares de pontos $[(\lambda_{min},4);(\lambda_{max},3)][(\lambda_{min}5);(\lambda_{max},3)]$ e $[(\lambda_{min},5);(\lambda_{max},4)]$.

O esquema geral para obter os parâmetros ótimos destes filmes é como se segue. Primeiro, dividimos o espectro em duas partes: $[\lambda_{min},\lambda_{bound}]$ e $[\lambda_{bound},\lambda_{max}]$, onde λ_{bound} é um limitante conhecido de λ_{infl} . Em todos os casos utilizamos $\lambda_{bound} = (\lambda_{min} + \lambda_{max})/2$. Para estimar a espessura utilizamos os dados com abcissa em $[\lambda_{bound},\lambda_{max}]$. O procedimento consiste em rodar o Algoritmo 3.1 para diferentes valores de d, entre $d_{min} = \frac{1}{2}d_{kick}$ e $d_{max} = \frac{3}{2}d_{kick}$ com passo $10\,(d_{min}, d_{min}+10, d_{min}+20, \ldots)$ onde d_{kick} pode ser uma estimativa inicial grosseira de d. Desta forma, obtemos d_{trial}, o valor da espessura que fornece o menor erro quadrático. Depois, repetimos o procedimento com $d_{min} = d_{trial} - 10, d_{max} = d_{trial} + 10$ e passo

1 para obter, finalmente, a espessura estimada d_{best}.

Para estimar o ponto de inflexão procedemos em forma análoga, utilizando o espectro completo, a espessura fixa em d_{best} e tentando diferentes pontos de inflexão (obviamente entre $\lambda_{min}, \lambda_{bound}$). Tomamos como estimativa de λ_{bound} o valor que nos fornece o menor erro quadrático. Em todas as simulações, permitimos só 3000 iterações do Algoritmo 3.1. O passo final do método consiste em, fixando d_{best} e λ_{infl}, rodar o Algoritmo 3.1 mais uma vez permitindo 30000 iterações.

Todos os experimentos foram simulados numa SPARCstation Sun Ultra 1, com um processador UltraSPARC de 64 bits, clock de 167 MHz e 128 Mbytes de memória RAM. Utilizamos a linguagem C/C++ com o compilador g++ (GNU project C and C++ compiler v2.7) e a opção de otimização do compilador -04. Apesar das muitas execuções do algoritmo de minimização irrestrita que são necessárias para resolver o problema, o tempo total de CPU utilizado para o processo completo nunca excedeu os 10 minutos na plataforma acima mencionada.

A Tabela 1 corresponde só ao filme A. Ela mostra a precisão obtida em $r(\lambda)$ e $\alpha(\lambda)$ utilizando 25, 50 e 100 pontos de transmissão medida, e arredondando os dados de transmissão para, 2, 3 e 4 casas decimais e, finalmente, sem arredondar. Os erros reportados são os máximos valores de $|r(\lambda) - r_{true}(\lambda)|$ e $|\alpha(\lambda) - \alpha_{true}(\lambda)|$

		25		50		100	
2	E_{ph-min}	r 0.1136	α 1.9021x10^{-4}	r 0.1036	α 4.3521x10^{-4}	r 0.1139	α 4.3108x10^{-4}
	E_{ph-max}	0.8742	4.4408x10^{-3}	1.4727	4.9223x10^{-3}	0.4761	1.9139x10^{-3}

Continuação da tabela 1.

3	E_{ph-min}	0.0442	$2.7240x10^{-4}$	0.1317	$6.5932x10^{-4}$	0.2249	$1.1136x10^{-3}$
	E_{ph-max}	1.3505	$4.4571x10^{-3}$	0.2298	$1.4237x10^{-3}$	0.1278	$1.0697x10^{-3}$
4	E_{ph-min}	0.0552	$3.3624x10^{-4}$	0.1093	$4.0392x10^{-4}$	0.1103	$4.3725x10^{-4}$
	E_{ph-max}	1.3631	$4.5240x10^{-3}$	0.1418	$1.0016x10^{-3}$	0.1149	$3.8185x10^{-4}$
todos	E_{ph-min}	0.0358	$3.9590x10^{-4}$	0.0749	$3.0955x10^{-4}$	0.0184	$1.6248x10^{-4}$
	E_{ph-max}	1.4558	$4.5031x10^{-3}$	0.2367	$7.5514x10^{-4}$	0.0117	$1.9621x10^{-4}$

Tabela 1: Filme A: Erros quadráticos nos índices de refração e coeficientes de absorção estimados variando a precisão e o número de pontos de transmissão medida utilizados.

Abaixo é representada a tabela 2: filme A

	25	50	100
2	121	121	121
3	119	122	124
4	119	121	121
todos	118	119	119

Tabela 2: Filme A: espessura estimada variando a precisão da transmissão e o número de pontos de transmissão medidos.

Para as regiões espectrais de alta e baixa energia do fóton, respectivamente. A Tabela 2 corresponde ao mesmo filme. Ela mostra a espessura estimada com 25, 50 e 100 pontos de dados e para diferentes números de casas decimais em $T^{meas}(\lambda)$. A Tabela 3 mostra, para todos os filmes, as espessuras estimadas e os erros quadráticos obtidos no processo de minimização.

A seguir, o objeto desta patente de invenção será pormenorizadamente descrito com referência aos desenhos abaixo relacionados, nos quais:

a figura 1 ilustra um conjunto de gráficos representativos dos valores "verdadeiros" e estimados da transmissão, o índice de refração e o coeficiente de absorção de um filme muito fino gerado numericamente com espessura *d* = 118nm que simula uma capa de a-Ge depositada sobre vidro, cabendo ressaltar a boa qualidade das constantes óticas e a transmissão encontradas;

a figura 2 ilustra por sua vez, três gráficos que são representativos dos valores "verdadeiros" e estimação da transmissão, o índice de refração e o coeficiente de absorção de um filme gerado numericamente com espessura *d* = 782nm simulando uma capa de a-Ge depositada sobre vidro, cabendo também notar a boa qualidade das constantes óticas e a transmissão encontradas;

a figura 3 ilustra também três gráficos que são representativos dos valores "verdadeiros" e estimados da transmissão, o índice de refração e o coeficiente de absorção de um filme muito fino gerado numericamente com espessura *d* = 147nm que simula uma capa de a-Ge depositada num substrato de silício cristalino, cabendo também neste caso ressaltar a qualidade das constantes óticas e a transmissão encontradas apesar da suavidade e falta de interferências da transmissão verdadeira;

a figura 4 ilustra um conjunto de gráficos que representam: valores "verdadeiros" e estimados da transmissão, o índice de refração e o coeficiente de absorção de um filme fino gerado numericamente com espessura d = 640nm simulando uma capa de a-Ge depositada num substrato de silício cristalino, notando-se também a qualidade das constantes óticas e a transmissão encontradas, sendo que, a recuperação do verdadeiro índice de refração mostra alguns defeitos na região de altas energias de fótons;

a figura 5 ilustra três gráficos que representam valores "verdadeiros" e estimados da transmissão, o índice de refração e o coeficiente de absorção de um filme fino gerado numericamente com espessura d = 624nm simulando uma capa de a-Si- depositada sobre vidro, sendo que, a coincidência das constantes óticas e da transmissão pode ser classificada como muito boa;

a figura 6 ilustra um conjunto de gráficos, representativos do erro quadrático do processo de minimização em função das espessuras para o filme A e B; Do lado esquerdo o passo é 10nm enquanto que do lado direito o passo refinado é 1nm, notando-se os minimizadores locais;

a figura 7 ilustra um conjunto de gráficos que representam o erro quadrático do processo de minimização em função das espessuras para o filme C e D, sendo que, do lado esquerdo o passo é 10nm enquanto do lado direito o passo refinado é 1nm, cabendo também neste caso a menção aos minimizadores locais; e

a figura 8 por sua vez, representa dois gráficos que tratam do erro quadrático do processo de minimização em função das espessuras para o filme E, sendo que, do lado esquerdo o passo é de 10nm enquanto que do lado direito o passo refinado é 1nm, cambendo também neste caso citar os minimizadores locais.

As figuras 1-5 são auto-expli

cáveis. As linhas contínuas representam as transmissões verdadeiras, os índices de refração verdadeiros e os coeficientes de absorção verdadeiros para os cinco filmes considerados. Os círculos abertos representam as melhores estimativas (utilizando todas as casas decimais). Notar que as funções $r(\lambda)$ e $\alpha(\lambda)$ estão representadas em função da energia do fóton $(1240/\lambda)$. Finalmente, as Figuras 6, 7 e 8 mostram como funciona o processo de estimar a espessura para cada filme.

A análise dos resultados numéricos permite-nos esboçar as seguintes conclusões: a) o procedimento proposto é altamente confiável para estimar a espessura verdadeira em todos os filmes quando 4 (ou todos) dígitos dos dados de transmissão são utilizados. O método fornece uma recuperação muito boa da transmissão verdadeira nos casos onde métodos não aproximados são úteis, *i.e.*, filmes muito finos ou camadas absorventes; b) apreciação dos dados da transmissão medida tem efeito na precisão das estimações de $r(\lambda)$ mas é quase irrelevante para a estimação de $\alpha(\lambda)$. Em situações reais, utilizando um espectrofotômetro moderno, as transmissões podem ser obtidas com 3 ou 4 casas decimais. Neste caso, e utilizando 100 pontos medidos de transmissão, o erro na estimação de $r(\lambda)$ é próximo de 0.11 (para o Filme A). As diferenças entre os resultados obtidos com 100 e 50 pontos não são significativas; c) na maioria dos casos, o erro quadrático em função da espessura estimada (Figuras 6, 7 e 8) é uma função com vários minimizadores locais não-globais. Portanto, a estratégia de separar a variável *d* das outras variáveis do problema de otimização parece ser correta, dado que tende a evitar convergências espúrias a estes minimizadores locais; e d) a comparação dos presentes resultados com os obtidos previamente, utilizando o algoritmo descrito por I.CHAMBOULEYRON, J.M. MARTÍNEZ, A.C.MORETTI e M.MULATO (1997 e 1998), parece confirmar que o novo método é, no mínimo, tão eficiente quanto o enfoque prévio de minimização com restrições. Além disso, o software resultante é mais portável e fácil de manipular.

Abaixo é representada a tabela 3, que trata das espessuras verdadeiras e estimadas e erros quadráticos.

Filme	Espessura verdadeira	Espessura estimada	Erro quadrático
A	118	119	$6.9296x10^{-7}$
B	782	782	$2.2031x10^{-7}$
C	147	152	$6.2249x10^{-6}$
D	640	639	$1.3653x10^{-6}$
E	624	624	$2.1210x10^{-7}$

Tabela 3: Espessuras verdadeiras e estimadas e erros quadráticos.

REIVINDICAÇÕES

1. "METODOLOGIA PARA ESTIMAÇÃO DAS CONSTANTES ÓPTICAS E DA ESPESSURA DE UM FILME FINO UTILIZANDO MINIMIZAÇÃO IRRESTRITA", do tipo que pode ser utilizada para determinar a estimação das constantes ópticas e espessura de filmes finos em geral, caracterizada pelo fato de que a formulação irrestrita do problema de estimação, onde a transmissão T de um filme fino absorvente depositado num substrato transparente grosso é dado por:

$$\mathrm{T} = \frac{Ax}{B - Cx + Dx^2}, \qquad (3)$$

onde: $A = 16s(r^2 + k^2)$ (4)

$$B = [(r+1)^2 + k^2)][(r+1)(r+s^2) + k^2\}, \qquad (5)$$

$$C = [(r^2 - 1 + k^2)(r^2 - s^2 + k^2) - 2k^2(s^2+1)]2\cos\varphi$$

$$-k[2(r^2 - s^2 + k^2) + (s^2+1)(r^2 - 1 + k^2)]2\,\mathrm{sen}\,\varphi\,, \qquad (6)$$

$$D = [(r-1)^2 + k^2][(r-1)(r-s^2) + k^2], \qquad (7)$$

$$\varphi = 4\pi rd/\lambda,\ x = \exp(-\alpha d),\ \alpha = 4\pi k/\lambda. \qquad (8)$$

sendo que nas fórmulas (4)-(8) é utilizada a seguinte notação:

(a) λ é o comprimento de onda;

(b) $s = s(\lambda)$ é o índice de refração do substrato transparente (que supomos conhecido);

(c) $r = r(\lambda)$ é o índice de refração do filme;

(d) $k = k(\lambda)$ é o índice de atenuação do filme (α é o coeficiente de absorção);

(e) d é a espessura do filme.

2. "METODOLOGIA PARA ESTIMAÇÃO DAS CONSTANTES ÓPTICAS E DA ESPESSURA DE UM FILME FINO UTILIZANDO MINIMIZAÇÃO IRRESTRITA", caracterizada pelo fato

de se utilizar de um algoritmo de minimização irrestrita, o qual compreende: Passo 1: Detectar se o ponto atual é estacionário; Se $\|g(x_k)\|=0$, parar declarando que x_k é estacionário;

Passo 2: Busca linear;

Passo 2.1: Fazer $\lambda \leftarrow \alpha_k$;

Passo 2.2: Fazer $x_+ = x_k - \lambda g(x_k)$;

Passo 2.3: Se $f(x_+) \leq max_{0\leq j\leq min\{k,M-1\}} f(x_{k-j}) + \gamma\langle x_+ - x_k, g(x_k)\rangle$, (38)

então definir $x_{k+1} = x_+, s_k = x_{k+1} - x_k$ e $y_k = g(x_{k+1}) - g(x_k)$;

Senão, definir $\lambda_{new} \in [\sigma_1\lambda, \sigma_2\lambda]$, (39)

Fazer $\lambda \leftarrow \lambda_{new}$ e voltar ao Passo 2.2;

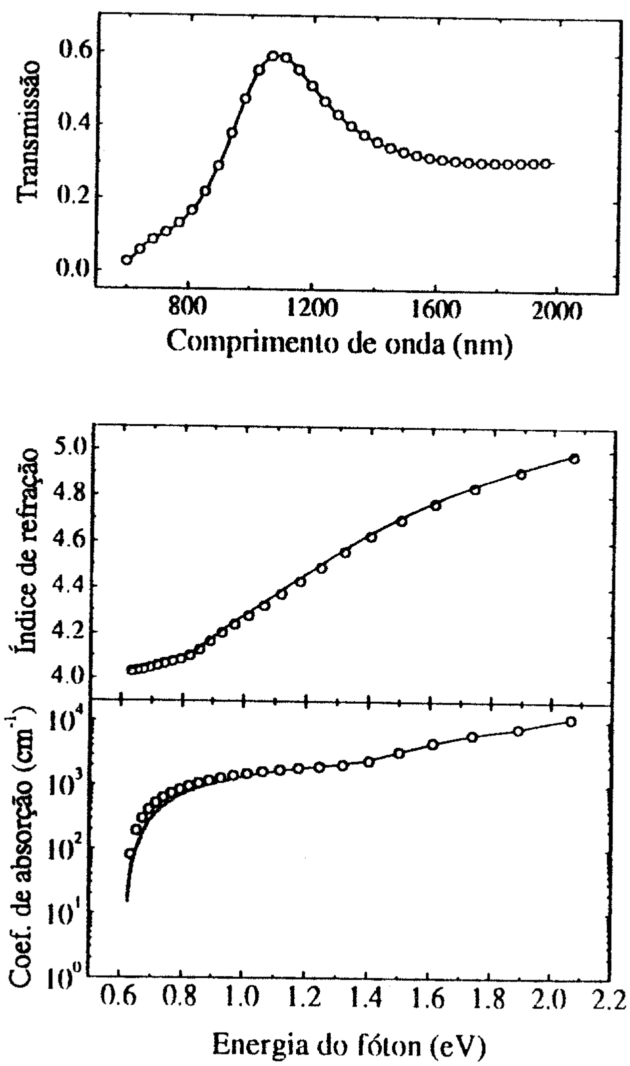
Passo 3: Calcular o comprimento de passo espectral;

Calcular $b_k = (s_k, y_k)$

Se $b_k \leq 0$, fazer $\alpha_{k+1} = \alpha_{max}$

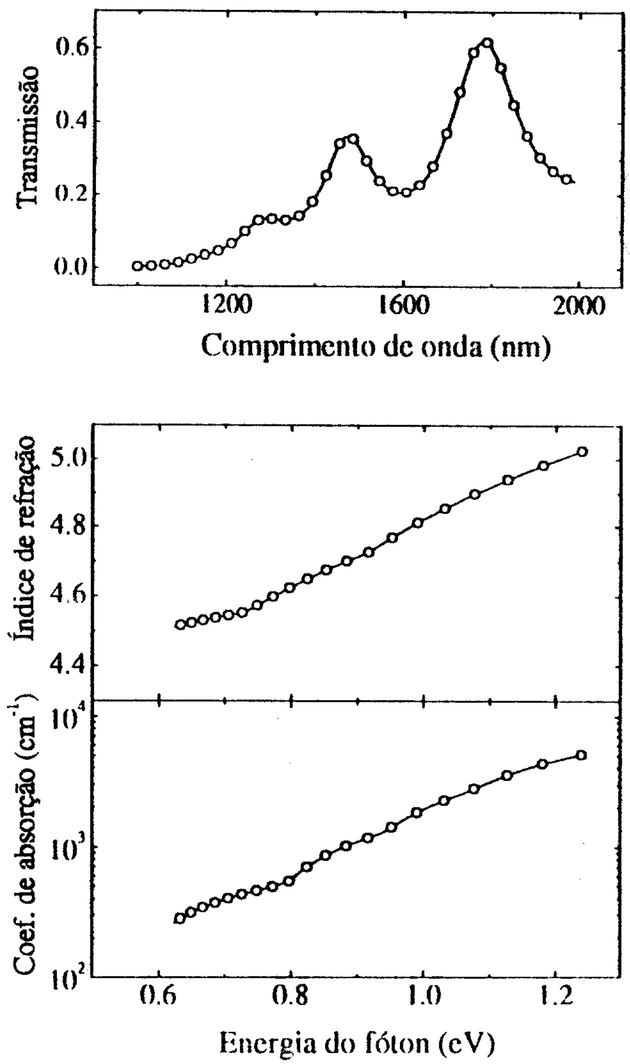
senão, calcular $a_k = -\langle s_k, g(x_k)\rangle$ e

$\alpha_{k+1} = \min\{\alpha_{max}, \max\{\alpha_{min}, a_k / b_k\}\}$



Transmissão
0.0
0.2
0.4
0.6
800
1200
1600
2000
Comprimento de onda (nm)
Índice de refração
4.0
4.2
4.4
4.6
4.8
5.0
Coef. de absorção (cm⁻¹)
10⁰
10¹
10²
10³
10⁴
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
Energia do fóton (eV)

FIG. -1



Transmissão
0.0
0.2
0.4
0.6
1200
1600
2000
Comprimento de onda (nm)
Índice de refração
4.4
4.6
4.8
5.0
Coef. de absorção (cm^{-1})
10^2
10^3
10^4
0.6
0.8
1.0
1.2
Energia do fóton (eV)

FIG. -2

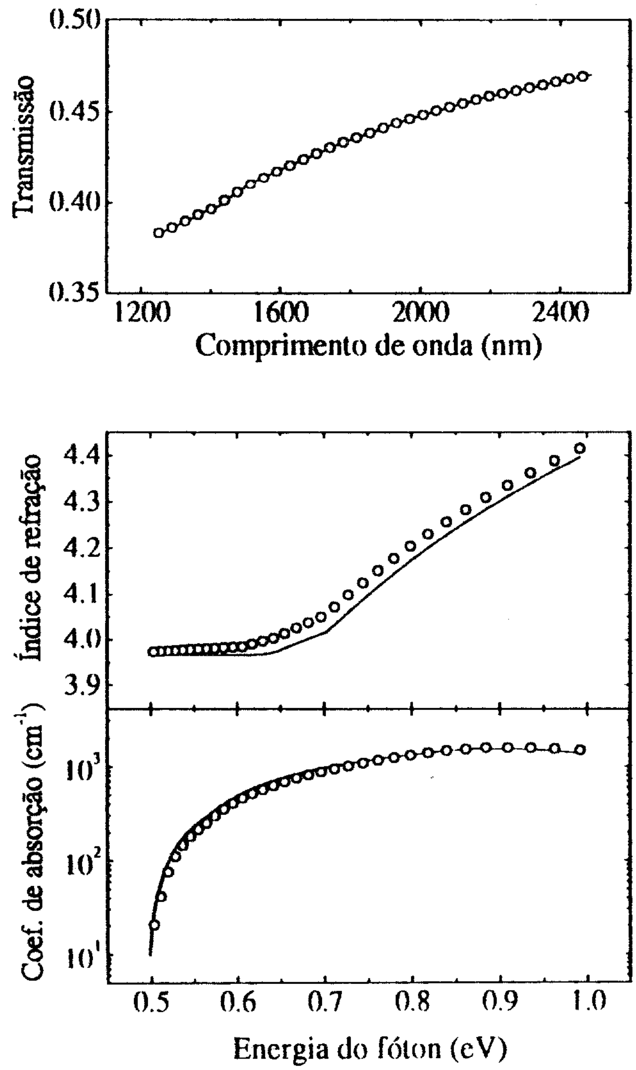

0.50
0.45
0.40
0.35
Transmissão
1200
1600
2000
2400
Comprimento de onda (nm)
4.4
4.3
4.2
4.1
4.0
3.9
Índice de refração
10^3
10^2
10^1
Coef. de absorção (cm^{-1})
0.5
0.6
0.7
0.8
0.9
1.0
Energia do fóton (eV)

FIG.-3

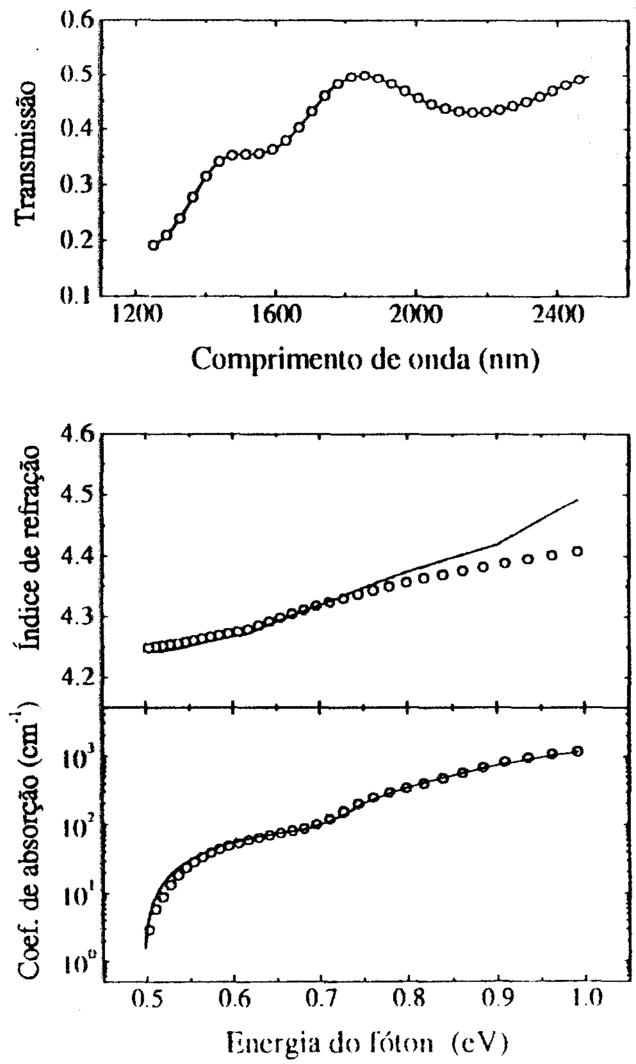
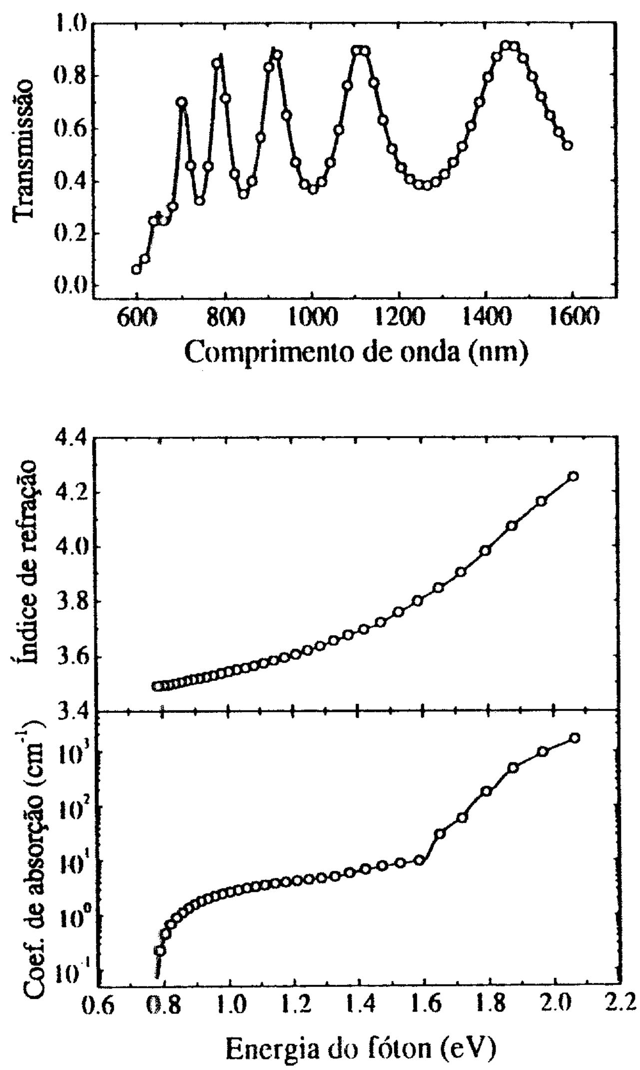
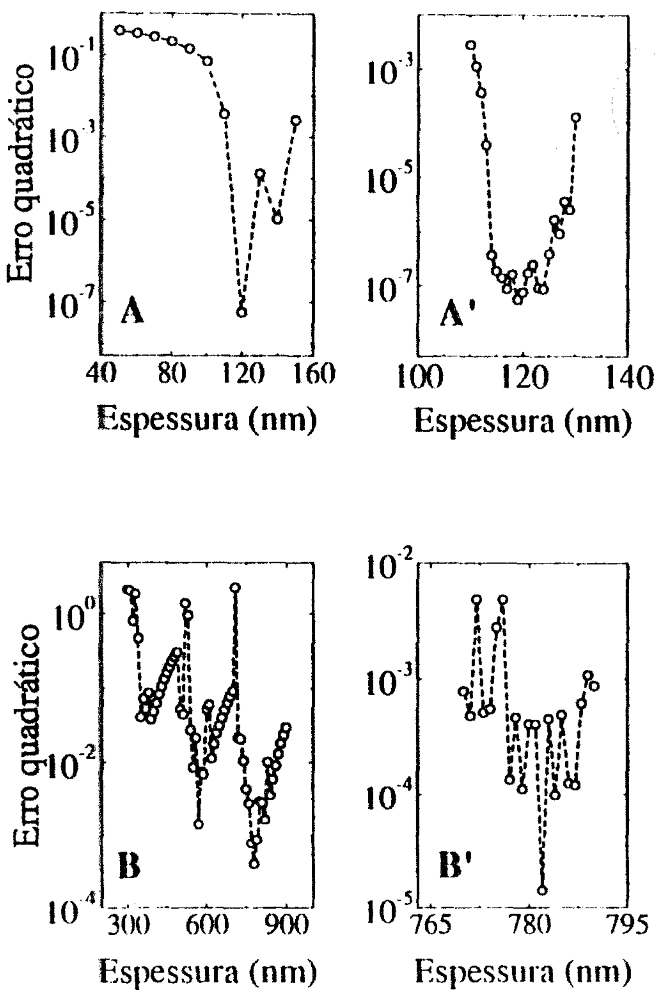

0.6
0.5
0.4
0.3
0.2
0.1
Transmissão
1200
1600
2000
2400
Comprimento de onda (nm)
4.6
4.5
4.4
4.3
4.2
Índice de refração
Coef. de absorção (cm-1)
0.5
0.6
0.7
0.8
0.9
1.0
Energia do fóton (eV)

FIG.-4



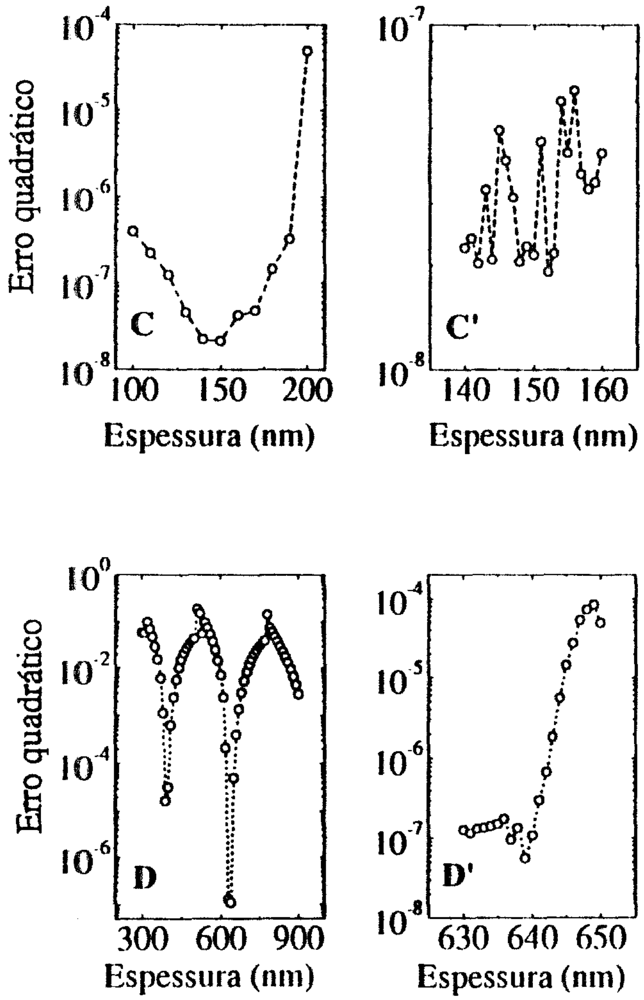
Transmissão
1.0
0.8
0.6
0.4
0.2
0.0
600
800
1000
1200
1400
1600
Comprimento de onda (nm)
Índice de refração
4.4
4.2
4.0
3.8
3.6
3.4
Coef. de absorção (cm^{-1})
10^{3}
10^{2}
10^{1}
10^{0}
10^{-1}
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
Energia do fóton (eV)

FIG.-5



Erro quadrático
10^{-1}
10^{-3}
10^{-5}
10^{-7}
A
40
80
120
160
Espessura (nm)
10^{-3}
10^{-5}
10^{-7}
A'
100
120
140
Espessura (nm)
Erro quadrático
10^{0}
10^{-2}
10^{-4}
B
300
600
900
Espessura (nm)
10^{-2}
10^{-3}
10^{-4}
10^{-5}
B'
765
780
795
Espessura (nm)

FIG.-6



Erro quadrático
10^-4
10^-5
10^-6
10^-7
10^-8
C
100 150 200
Espessura (nm)
10^-7
10^-8
C'
140 150 160
Espessura (nm)
Erro quadrático
10^0
10^-2
10^-4
10^-6
D
300 600 900
Espessura (nm)
10^-4
10^-5
10^-6
10^-7
10^-8
D'
630 640 650
Espessura (nm)

FIG-7

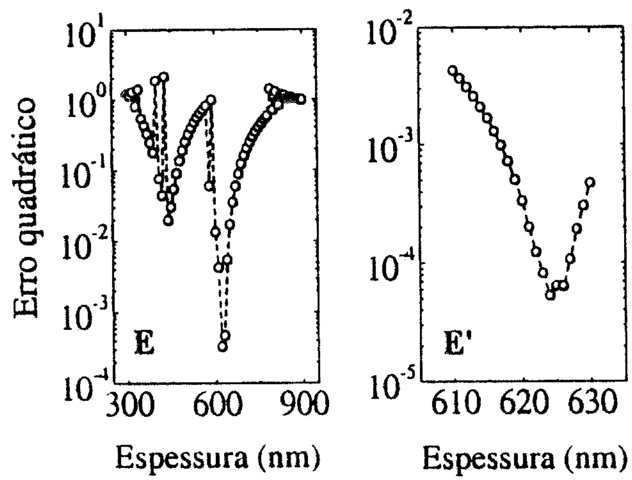

Erro quadrático
E
E'
10^1
10^0
10^-1
10^-2
10^-3
10^-4
10^-2
10^-3
10^-4
10^-5
300 600 900
610 620 630
Espessura (nm)
Espessura (nm)

FIG.-8

RESUMO

"METODOLOGIA PARA ESTIMAÇÃO DAS CONSTANTES ÓPTICAS E DA ESPESSURA DE UM FILME FINO UTILIZANDO MINIMIZAÇÃO IRRESTRITA", dita metodologia destina-se a possibilitar a estimação das constantes ópticas e da espessura de um filme fino, baseada na minimização irrestrita.